Zero-determinant strategies under observation errors in repeated games

Azumi Mamiya and Genki Ichinose*
Department of Mathematical and Systems Engineering, Shizuoka University, Hamamatsu 432-8561, Japan
(Dated: July 2020)

Zero-determinant (ZD) strategies are a novel class of strategies in the repeated prisoner's dilemma (RPD) game discovered by Press and Dyson. This strategy set enforces a linear payoff relationship between a focal player and the opponent regardless of the opponent's strategy. In the RPD game, games with discounting and observation errors represent an important generalization, because they are better able to capture real life interactions which are often noisy. However, they have not been considered in the original discovery of ZD strategies. In some preceding studies, each of them has been considered independently. Here, we analytically study the strategies that enforce linear payoff relationships in the RPD game considering both a discount factor and observation errors. As a result, we first reveal that the payoffs of two players can be represented by the form of determinants as shown by Press and Dyson even with the two factors. Then, we search for all possible strategies that enforce linear payoff relationships and find that both ZD strategies and unconditional strategies are the only strategy sets to satisfy the condition. We also show that neither Extortion nor Generous strategies, which are subsets of ZD strategies, exist when there are errors. Finally, we numerically derive the threshold values above which the subsets of ZD strategies exist. These results contribute to a deep understanding of ZD strategies in society.

I. INTRODUCTION

Cooperation is a basis for building sustainable societies. In a one-shot interaction, cooperation among individuals is suppressed because cooperation takes costs to the actor while defection does not. This cooperation-defection relationship is well captured by the prisoner's dilemma (PD) game utilized in game theory. In the one-shot PD game, defection is the only Nash equilibrium. When the game is repeated, the situation drastically changes, which is modeled by the repeated prisoner's dilemma (RPD) game [1]. In the RPD game, cooperation will be rewarded by the opponent in the future. In such a situation, cooperation becomes a possible equilibrium. This mechanism is called direct reciprocity [2–4] and makes it possible for players to mutually cooperate in the RPD game.

Evolutionary game theory (EGT) [5] studies how cooperation evolves in the RPD game. Among various cooperative strategies tested in evolutionary games, generous tit-for-tat [6] and win-stay lose-shift [7, 8] were robust to various kinds of evolutionary opponents under noisy conditions. EGT can find strong strategies against various opponents in evolving populations. One missing point was, what is a strong strategy against a direct opponent which utilizes any kind of strategy? In 2012, Press and Dyson suddenly answered this question from a different point of view. Using linear algebraic manipulations, they found a novel class of strategies which contain such ultimate strategies, called zero-determinant (ZD) strategies [9]. ZD strategies impose a linear relationship between the payoffs for a focal player and his opponent regardless of the strategy that the opponent implements. One of the subclasses of ZD strategies is Extortioner which never loses in a one-to-one competition in the RPD game against any opponents.

The discovery of ZD strategies stimulated many researchers. After Stewart and Plotkin raised a question [10], evolution or emergence of ZD strategies became one of the main targets in subsequent studies [11–25]. Then, this research spread in many directions including multiplayer games [19, 26–29], continuous action spaces [28–31], alternating games [31], asymmetric games [32], animal contests [33], human reactions to computerized ZD strategies [34, 35], and human-human experiments [28, 36, 37], which promote an understanding of the nature of human cooperation. For further understanding, see the recent elegant classification of strategies, partners (called "good strategies" in Ref. [11, 38]) and rivals, in direct reciprocity [39]. The utilization of ZD strategies has recently expanded to engineering fields, not just for human cooperation [40–42].

In those ZD studies, no errors were assumed. However, errors (or noise) are unavoidable in human interactions and they may lead to the collapse of cooperation due to negative effects. Thus, the effect of errors has been focused on in the RPD game [43–51]. However, only a few studies have concerned the effect of errors for ZD strategies [52, 53]. There are typically two types of errors: perception errors [45] and implementation errors [46]. Hao et al. [52] and Mamiya and Ichinose [53] considered the former case of the errors where players may misunderstand their opponent's action because the players can only rely on their private monitoring [43, 47] instead of their opponent's direct action. Those studies showed that ZD strategies can exist even in the case

* ichinose.genki@shizuoka.ac.jp;

that such observation errors are incorporated. In those studies, no discount factor is considered. It is natural to assume that future payoffs will be discounted. Thus, some studies have focused on a discount factor for ZD strategies [30, 31, 54–56] and mathematically found the minimum discount factor above which the ZD strategies can exist [55].

In this study, we search for ZD strategies under the situations that observation errors and a discount factor are both incorporated. We search for the other possible strategies, not just ZD strategies, that enforce a linear payoff relationship between the two players. By formalizing the determinants for the expected payoffs in the RPD game, we mathematically found that only ZD strategies [9] and unconditional strategies [14, 55] are the two types which enforce a linear payoff relationship. We numerically show the threshold values above which the subsets of ZD strategies exist in the game.

II. MODEL

II.1. RPD with private monitoring

We consider the symmetric two-person repeated prisoner's dilemma (RPD) game with private monitoring based on the literature [47, 52]. Each player $i \in \{X, Y\}$ chooses an action $a_i \in \{\mathrm{C}, \mathrm{D}\}$ in each round, where C and D imply cooperation and defection, respectively. After the two players conduct the action, player i observes his own action a_i and private signal $\omega_i \in \{g, b\}$ about the opponent's action, where g and b imply *good* and *bad*, respectively. In perfect monitoring, when the opponent takes the action C (D), the focal player always observes the signal g (b). In private monitoring, this is not always true. $\sigma(\boldsymbol{\omega}|\boldsymbol{a})$ is the probability that a signal profile $\boldsymbol{\omega} = (\omega_X, \omega_Y)$ is realized when the action profile is $\boldsymbol{a} = (a_X, a_Y)$ [47]. Let ϵ be the probability that an error occurs to one particular player but not to the other player while ξ be the probability that an error occurs to both players. Then, the probability that an error occurs to neither player is $1 - 2\epsilon - \xi$. For example, when both players take cooperation, $\sigma((g, g)|(\mathrm{C}, \mathrm{C})) = 1 - 2\epsilon - \xi$, $\sigma((g, b)|(\mathrm{C}, \mathrm{C})) = \sigma((b, g)|(\mathrm{C}, \mathrm{C})) = \epsilon$, and $\sigma((b, b)|(\mathrm{C}, \mathrm{C})) = \xi$ are realized.

In each round, player i's realized payoff $u_i(a_i, \omega_i)$ is determined by his own action a_i and signal ω_i, such that $u_i(\mathrm{C}, g) = R$, $u_i(\mathrm{C}, b) = S$, $u_i(\mathrm{D}, g) = T$, and $u_i(\mathrm{D}, b) = P$. Note that the payoffs depend on the signals in private monitoring. Hence, his expected payoff is given by

$$f_i(\boldsymbol{a}) = \sum_{\omega} u_i(a_i, \omega_i)\sigma(\boldsymbol{\omega}|\boldsymbol{a}). \tag{1}$$

The expected payoff is determined by only action profile $\boldsymbol{a}$ regardless of signal profile $\boldsymbol{\omega}$. Thus, the expected payoff matrix is given by

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{C} & \mathrm{D} \end{array} \\ \begin{array}{c} \mathrm{C} \\ \mathrm{D} \end{array} & \begin{pmatrix} R_E & S_E \\ T_E & P_E \end{pmatrix} \end{array}. \tag{2}$$

According to Eq. (1), R_E, S_E, T_E, and P_E are derived as $R_E = R(1-\epsilon-\xi)+S(\epsilon+\xi)$, $S_E = S(1-\epsilon-\xi)+R(\epsilon+\xi)$, $T_E = T(1-\epsilon-\xi)+P(\epsilon+\xi)$, $P_E = P(1-\epsilon-\xi)+T(\epsilon+\xi)$, respectively. We assume that

$$T_E > R_E > P_E > S_E, \tag{3}$$

and

$$2R_E > T_E + S_E, \tag{4}$$

which dictate the RPD condition with observation errors.

In this paper, we introduce a discount factor to the RPD game with private monitoring. The game is to be played repeatedly over an infinite time horizon but the payoff will be discounted over rounds. Player i's discounted payoff of action profiles $\boldsymbol{a}(t)$, $t \in \{0, 1, ..., \infty\}$ is $\delta^t f_i(\boldsymbol{a}(t))$ where δ is a discount factor and t is a round. This game can be interpreted as repeated games with a finite but undetermined time horizon. Finally, the average discounted payoff of player i is

$$s_i = (1 - \delta)\sum_{t=0}^{\infty} \delta^t f_i(\boldsymbol{a}(t)). \tag{5}$$

II.2. Determinant form of expected payoff in the RPG game

Here, we proceed to show that Eq. (5) can be represented by a determinant form even for the repeated games with observation errors and a discount factor, as Press and Dyson did for the repeated game without error and no discount factor [9]. The action profiles $\boldsymbol{a}(t)$ in Eq. (5) need to be specified to calculate s_i. Those profiles are determined after the strategies of two players are given. Consider player i that adopts memory-one strategies, with which they can use only the outcomes of the last round to decide the action to be submitted in the current round. A memory-one strategy is specified by a 5-tuple; X's strategy is given by a combination of

$$\boldsymbol{p} = (p_1, p_2, p_3, p_4; p_0), \tag{6}$$

where $0 \le p_j \le 1, j \in \{0, 1, 2, 3, 4\}$. The subscripts 1, 2, 3, and 4 of p mean previous outcomes Cg, Cb, Dg, and Db, respectively. In Eq. (6), p_1 is the conditional probability that X cooperates when X cooperated and observed signal g in the last round, p_2 is the conditional probability that X cooperates when X cooperated and observed signal b in the last round, p_3 is the conditional probability

that X cooperates when X defected and observed signal g in the last round, and p_4 is the conditional probability that X cooperates when X defected and observed signal b in the last round. Finally, p_0 is the probability that X cooperates in the first round. Similarly, Y's strategy is specified by a combination of

$$\boldsymbol{q} = (q_1, q_2, q_3, q_4; q_0), \tag{7}$$

where $0 \leq q_j \leq 1, j \in \{0, 1, 2, 3, 4\}$.

Define $\boldsymbol{v}(t) = (v_1(t), v_2(t), v_3(t), v_4(t))$ as the stochastic state of two players in round t where the subscripts 1, 2, 3, and 4 of v imply the stochastic states (C,C), (C,D), (D,C), and (D,D), respectively. $v_1(t)$ is the probability that both players cooperate in round t, $v_2(t)$ is the probability that X cooperates and Y defects in round t, and so forth. Then, the expected payoff to player X in round t is given by $\boldsymbol{v}(t)\boldsymbol{S}_X$, where $\boldsymbol{S}_X^T = (R_E, S_E, T_E, P_E)$. The expected per-round payoff to player X in the repeated game is given by

$$s_X = (1-\delta)\sum_{t=0}^{\infty}\delta^t \boldsymbol{v}(t)\boldsymbol{S}_X, \tag{8}$$

where $0 < \delta < 1$. The initial stochastic state is given by

$$\boldsymbol{v}(0) = (p_0 q_0, p_0(1-q_0), (1-p_0)q_0, (1-p_0)(1-q_0)). \tag{9}$$

The state transition matrix M of these repeated games with observation errors is given by

$$M = \begin{pmatrix} \begin{pmatrix} \tau p_1 q_1 \\ +\epsilon p_1 q_2 \\ +\epsilon p_2 q_1 \\ +\xi p_2 q_2 \end{pmatrix} & \begin{pmatrix} \tau p_1(1-q_1) \\ +\epsilon p_1(1-q_2) \\ +\epsilon p_2(1-q_1) \\ +\xi p_2(1-q_2) \end{pmatrix} & \begin{pmatrix} \tau(1-p_1)q_1 \\ +\epsilon(1-p_1)q_2 \\ +\epsilon(1-p_2)q_1 \\ +\xi(1-p_2)q_2 \end{pmatrix} & \begin{pmatrix} \tau(1-p_1)(1-q_1) \\ +\epsilon(1-p_1)(1-q_2) \\ +\epsilon(1-p_2)(1-q_1) \\ +\xi(1-p_2)(1-q_2) \end{pmatrix} \\ \begin{pmatrix} \epsilon p_1 q_3 \\ +\xi p_1 q_4 \\ +\tau p_2 q_3 \\ +\epsilon p_2 q_4 \end{pmatrix} & \begin{pmatrix} \epsilon p_1(1-q_3) \\ +\xi p_1(1-q_4) \\ +\tau p_2(1-q_3) \\ +\epsilon p_2(1-q_4) \end{pmatrix} & \begin{pmatrix} \epsilon(1-p_1)q_3 \\ +\xi(1-p_1)q_4 \\ +\tau(1-p_2)q_3 \\ +\epsilon(1-p_2)q_4 \end{pmatrix} & \begin{pmatrix} \epsilon(1-p_1)(1-q_3) \\ +\xi(1-p_1)(1-q_4) \\ +\tau(1-p_2)(1-q_3) \\ +\epsilon(1-p_2)(1-q_4) \end{pmatrix} \\ \begin{pmatrix} \epsilon p_3 q_1 \\ +\tau p_3 q_2 \\ +\xi p_4 q_1 \\ +\epsilon p_4 q_2 \end{pmatrix} & \begin{pmatrix} \epsilon p_3(1-q_1) \\ +\tau p_3(1-q_2) \\ +\xi p_4(1-q_1) \\ +\epsilon p_4(1-q_2) \end{pmatrix} & \begin{pmatrix} \epsilon(1-p_3)q_1 \\ +\tau(1-p_3)q_2 \\ +\xi(1-p_4)q_1 \\ +\epsilon(1-p_4)q_2 \end{pmatrix} & \begin{pmatrix} \epsilon(1-p_3)(1-q_1) \\ +\tau(1-p_3)(1-q_2) \\ +\xi(1-p_4)(1-q_1) \\ +\epsilon(1-p_4)(1-q_2) \end{pmatrix} \\ \begin{pmatrix} \xi p_3 q_3 \\ +\epsilon p_3 q_4 \\ +\epsilon p_4 q_3 \\ +\tau p_4 q_4 \end{pmatrix} & \begin{pmatrix} \xi p_3(1-q_3) \\ +\epsilon p_3(1-q_4) \\ +\epsilon p_4(1-q_3) \\ +\tau p_4(1-q_4) \end{pmatrix} & \begin{pmatrix} \xi(1-p_3)q_3 \\ +\epsilon(1-p_3)q_4 \\ +\epsilon(1-p_4)q_3 \\ +\tau(1-p_4)q_4 \end{pmatrix} & \begin{pmatrix} \xi(1-p_3)(1-q_3) \\ +\epsilon(1-p_3)(1-q_4) \\ +\epsilon(1-p_4)(1-q_3) \\ +\tau(1-p_4)(1-q_4) \end{pmatrix} \end{pmatrix}, \tag{10}$$

where $\tau = 1 - 2\epsilon - \xi$. Then, we obtain

$$\boldsymbol{v}(t) = \boldsymbol{v}(0)M^t. \tag{11}$$

By substituting Eq. (11) in Eq (8), we obtain

$$\begin{aligned} s_X &= (1-\delta)\boldsymbol{v}(0)\sum_{t=0}^{\infty}(\delta M)^t \boldsymbol{S}_X \\ &= (1-\delta)\boldsymbol{v}(0)(I-\delta M)^{-1}\boldsymbol{S}_X, \end{aligned} \tag{12}$$

where I is the 4×4 identity matrix. Then, let

$$\boldsymbol{v}^T = (v_1, v_2, v_3, v_4) = (1-\delta)\boldsymbol{v}(0)(I-\delta M)^{-1} \tag{13}$$

be the mean distribution of $\boldsymbol{v}(t)$. Additionally, we define

$$M_0 = \begin{pmatrix} p_0 q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0 q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0 q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0 q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \end{pmatrix}. \tag{14}$$

Because $v_1 + v_2 + v_3 + v_4 = 1$ (Appendix A), the following holds (Appendix B):

$$\boldsymbol{v}(0) = \boldsymbol{v}^T M_0. \tag{15}$$

By substituting Eq. (15) in Eq. (13) and multiplying both sides of the equation by $(I - \delta M)$ from the right, we obtain

$$\boldsymbol{v}^T(I-\delta M) = (1-\delta)\boldsymbol{v}^T M_0. \tag{16}$$

Equation (16) and $M' \equiv \delta M + (1-\delta)M_0 - I$ yield

$$\boldsymbol{v}^T M' = 0. \tag{17}$$

With Eq. (17), we immediately obtain a formula for the dot product of an arbitrary vector $\boldsymbol{f}^T = (f_1, f_2, f_3, f_4)$ with the fourth column vector $\boldsymbol{u}$ of matrix M' as a consequence of Press and Dyson's formalism, which can be represented by the form of the determinant

$$\boldsymbol{u}\cdot\boldsymbol{f} = \begin{vmatrix} \delta(\tau p_1q_1+\epsilon p_1q_2+\epsilon p_2q_1+\xi p_2q_2)-1+p_0q_0(1-\delta) & \delta(\mu p_1+\eta p_2)-1+p_0(1-\delta) & \delta(\mu q_1+\eta q_2)-1+q_0(1-\delta) & f_1 \\ \delta(\epsilon p_1q_3+\xi p_1q_4+\tau p_2q_3+\epsilon p_2q_4)+p_0q_0(1-\delta) & \delta(\eta p_1+\mu p_2)-1+p_0(1-\delta) & \delta(\mu q_3+\eta q_4)+q_0(1-\delta) & f_2 \\ \delta(\epsilon p_3q_1+\tau p_3q_2+\xi p_4q_1+\epsilon p_4q_2)+p_0q_0(1-\delta) & \delta(\mu p_3+\eta p_4)+p_0(1-\delta) & \delta(\eta q_1+\mu q_2)-1+q_0(1-\delta) & f_3 \\ \delta(\xi p_3q_3+\epsilon p_3q_4+\epsilon p_4q_3+\tau p_4q_4)+p_0q_0(1-\delta) & \delta(\eta p_3+\mu p_4)+p_0(1-\delta) & \delta(\eta q_3+\mu q_4)+q_0(1-\delta) & f_4 \end{vmatrix} \equiv D(\boldsymbol{p},\boldsymbol{q},\boldsymbol{f}), \tag{18}$$

where $\mu = 1-\epsilon-\xi$ and $\eta = \epsilon+\xi$. Furthermore, Eq. (18) should be normalized to have its components sum to 1 by $\boldsymbol{u}\cdot\mathbf{1}$, where $\mathbf{1} = (1,1,1,1)$. Then, we obtain the dot product of an arbitrary vector $\boldsymbol{f}$ with mean distribution $\boldsymbol{v}$. Replacing the last column of $D(\boldsymbol{p},\boldsymbol{q},\boldsymbol{f})$ with player X's and Y's expected payoff vector, respectively, we obtain their per-round expected payoffs:

$$s_X = \boldsymbol{v}\cdot\boldsymbol{S}_X = \frac{\boldsymbol{u}\cdot\boldsymbol{S}_X}{\boldsymbol{u}\cdot\mathbf{1}} = \frac{D(\boldsymbol{p},\boldsymbol{q},\boldsymbol{S}_X)}{D(\boldsymbol{p},\boldsymbol{q},\mathbf{1})}, \tag{19}$$

$$s_Y = \boldsymbol{v}\cdot\boldsymbol{S}_Y = \frac{\boldsymbol{u}\cdot\boldsymbol{S}_Y}{\boldsymbol{u}\cdot\mathbf{1}} = \frac{D(\boldsymbol{p},\boldsymbol{q},\boldsymbol{S}_Y)}{D(\boldsymbol{p},\boldsymbol{q},\mathbf{1})}. \tag{20}$$

When we set $\delta = 1$, Eq. (18) corresponds to Eq. (2) of [52]. By using Eq. (18), we can calculate players' per-round expected payoffs when $0 < \delta \leq 1$ by the form of the determinants. $\delta = 1$ is the case where future payoffs are not discounted.

III. RESULTS

III.1. All strategies that enforce linear payoff relationships

Since we are interested in the payoff relationship between the two players, we linearly combine those payoffs represented by Eqs. (19) and (20). The linear combination of s_X and s_Y can also be represented by the form of the determinant:

$$\alpha s_X + \beta s_Y + \gamma = \frac{D(\boldsymbol{p},\boldsymbol{q},\alpha\boldsymbol{S}_X+\beta\boldsymbol{S}_Y+\gamma\mathbf{1})}{D(\boldsymbol{p},\boldsymbol{q},\mathbf{1})}, \tag{21}$$

where α, β, and γ, are arbitrary constants. The numerator of the right side of Eq. (21) is expressed in the following way:

$$D(\boldsymbol{p},\boldsymbol{q},\alpha\boldsymbol{S}_X+\beta\boldsymbol{S}_Y+\gamma\mathbf{1}) = \begin{vmatrix} \delta(\tau p_1q_1+\epsilon p_1q_2+\epsilon p_2q_1+\xi p_2q_2)-1+p_0q_0(1-\delta) & \delta(\mu p_1+\eta p_2)-1+p_0(1-\delta) & \delta(\mu q_1+\eta q_2)-1+q_0(1-\delta) & \alpha R_E+\beta R_E+\gamma \\ \delta(\epsilon p_1q_3+\xi p_1q_4+\tau p_2q_3+\epsilon p_2q_4)+p_0q_0(1-\delta) & \delta(\eta p_1+\mu p_2)-1+p_0(1-\delta) & \delta(\mu q_3+\eta q_4)+q_0(1-\delta) & \alpha S_E+\beta T_E+\gamma \\ \delta(\epsilon p_3q_1+\tau p_3q_2+\xi p_4q_1+\epsilon p_4q_2)+p_0q_0(1-\delta) & \delta(\mu p_3+\eta p_4)+p_0(1-\delta) & \delta(\eta q_1+\mu q_2)-1+q_0(1-\delta) & \alpha T_E+\beta S_E+\gamma \\ \delta(\xi p_3q_3+\epsilon p_3q_4+\epsilon p_4q_3+\tau p_4q_4)+p_0q_0(1-\delta) & \delta(\eta p_3+\mu p_4)+p_0(1-\delta) & \delta(\eta q_3+\mu q_4)+q_0(1-\delta) & \alpha P_E+\beta P_E+\gamma \end{vmatrix}. \tag{22}$$

If Eq. (22) is zero, the relationship between the two players' payoffs becomes linear:

$$\alpha s_X + \beta s_Y + \gamma = 0. \tag{23}$$

Thus, we search for all of the solutions such that $D(\boldsymbol{p},\boldsymbol{q},\alpha\boldsymbol{S}_X+\beta\boldsymbol{S}_Y+\gamma\mathbf{1}) = 0$.

Press and Dyson [9] (without error) and Hao et al. [52] (with observation errors) searched for the case that second and fourth columns of the determinant take the same value. This makes the determinant zero. Also, Mamiya and Ichinose [53] searched for all the cases, from all possibilities, that make the determinant zero with observation errors. Here, we extend Mamiya and Ichinose [53] to the case with both observation errors and a discount factor.

The following determinant theorem gives such a condition:

$$det(\mathrm{A}) = 0 \Leftrightarrow \text{The columns of matrix A are linearly dependent vectors}$$

for $n\times n$ matrix A. We define $\boldsymbol{d}_i, i \in \{1,2,3,4\}$ as i-th column vector of the determinant of Eq. (22). From the above theorem, if the columns of the determinant of Eq. (22) are linearly dependent vectors, there exist real numbers s,t,u,v,α,β, and γ, except for the trivial solution $((s,t,u,v) = (0,0,0,0),(\alpha,\beta,\gamma) = (0,0,0))$, such that

$$s\boldsymbol{d}_1 + t\boldsymbol{d}_2 + u\boldsymbol{d}_3 + v\boldsymbol{d}_4 = \mathbf{0}, \tag{24}$$

where vector $\mathbf{0}$ denotes a zero vector. We give the detailed calculation in Appendix C.

As a result, we found that, in the RPD game even with observation errors (imperfect monitoring) and a discount factor, the only strategies that impose a linear payoff relationship between the two players' payoffs are either

$$\begin{aligned}\delta(\mu p_1+\eta p_2)-1+p_0(1-\delta)&=\alpha R_E+\beta R_E+\gamma,\\ \delta(\eta p_1+\mu p_2)-1+p_0(1-\delta)&=\alpha S_E+\beta T_E+\gamma,\\ \delta(\mu p_3+\eta p_4)+p_0(1-\delta)&=\alpha T_E+\beta S_E+\gamma,\\ \delta(\eta p_3+\mu p_4)+p_0(1-\delta)&=\alpha P_E+\beta P_E+\gamma\end{aligned} \tag{25}$$

or

$$p_0=p_1=p_2=p_3=p_4. \tag{26}$$

The former corresponds to ZD strategies and the latter corresponds to unconditional strategies, respectively.

III.2. Extortion and Generous no longer exist when there are errors

Extortion [9] and Generous [23] strategies are well-known subsets of ZD strategies, which have important characteristics. Extortion never loses to any opponent in a one-to-one competition in terms of the expected payoffs. Moreover, they finally impose ALLC (always cooperate) to the opponent who tries to improve his payoff [9]. On the other hand, Generous strategies always obtain lower payoffs than the opponent except for mutual cooperation. Hence, Generous strategies are known as one of the cooperative ZD strategies. Generous strategies are weak in a one-to-one competition. However, in a large evolving population, cooperative groups are more successful than the group of Extortioners. Thus, evolution leads from Extortion to Generous strategies.

In Eq. (25), we substitute $\alpha=\phi, \beta=-\phi\chi$, and $\gamma=\phi(\chi-1)\kappa$ [55] to obtain

$$\begin{aligned}\delta(\mu p_1+\eta p_2)-1+p_0(1-\delta)&=\phi[(R_E-\kappa)-\chi(R_E-\kappa)],\\ \delta(\eta p_1+\mu p_2)-1+p_0(1-\delta)&=\phi[(S_E-\kappa)-\chi(T_E-\kappa)],\\ \delta(\mu p_3+\eta p_4)+p_0(1-\delta)&=\phi[(T_E-\kappa)-\chi(S_E-\kappa)],\\ \delta(\eta p_3+\mu p_4)+p_0(1-\delta)&=\phi[(P_E-\kappa)-\chi(P_E-\kappa)],\end{aligned} \tag{27}$$

where $\kappa=P_E$ with $1\leq\chi<\infty$ represents Extortion while $\kappa=R_E$ with $1\leq\chi<\infty$ represents Generous strategies. The parameter χ gives the correlation between two players' payoffs. Thus, we call χ a correlation factor. The parameter κ corresponds to the payoff that the ZD strategy would obtain against itself. We thus call κ baseline payoff as used in Ref. [14].

Those two prominent strategy sets exist when there are no errors. Here, by contrast, we prove that Extortion and Generous strategies no longer exist when there are errors. The detail calculation is provided in Appendix D.

However, Hao et al. [52] and Mamiya and Ichinose [53] found that there exist ZD strategies which partially have the characteristic of Extortion even when there are errors. In the original meaning, Extortion has two properties: (1) his payoff increase is always higher than any opponent's due to $\chi>1$ (payoff control ability) and (2) he is not outperformed by anyone (payoff dominance). When there are errors, the second property is lost although the first property remains. Intuitively, this is because errors introduce uncertainty into the payoffs and, consequently, enforce a negative impact on the accuracy of player X's payoff-based strategy setting. To keep the payoff control ability, the payoff dominance needs to be sacrificed when there are errors. Hao et al. called the ZD strategies which only have the first property *contingent extortion* [52].

III.3. Existence of subsets of ZD strategies

Since observation errors and a discount factor are considered, in general, the ranges in which ZD strategies can exist are narrowed. Ichinose and Masuda mathematically showed the minimum threshold values above which Equalizer (another subclass of ZD strategies), Extortion, and Generous strategies can exist [55]. Here, we numerically address threshold values where subsets of ZD strategies can exist.

Minimum discount factor for Equalizer

Equalizer strategies are a subclass of ZD strategies. They can fix the expected payoffs of the opponent regardless of the opponent's strategies [9]. We first show minimum discount factor δ_c for Equalizer when observation errors ϵ and ξ are given. Equalizer can fix the opponent payoff no matter what the opponent takes, which means that

$$\beta s_Y+\gamma=0. \tag{28}$$

This is obtained by substituting $\alpha=0$ into Eq. (23). Note that $\chi\to\infty$ in Eq. (27) corresponds to Equalizer [52]. We substitute $\alpha=0$ into Eq. (25) to obtain Equalizer:

$$\begin{aligned}\delta(\mu p_1+\eta p_2)-1+p_0(1-\delta)&=\beta R_E+\gamma,\\ \delta(\eta p_1+\mu p_2)-1+p_0(1-\delta)&=\beta T_E+\gamma,\\ \delta(\mu p_3+\eta p_4)+p_0(1-\delta)&=\beta S_E+\gamma,\\ \delta(\eta p_3+\mu p_4)+p_0(1-\delta)&=\beta P_E+\gamma.\end{aligned} \tag{29}$$

If we solve Eq. (29) for β,γ,p_2 and p_3,

$$
\begin{aligned}
\beta &= -\frac{(1-\delta p_1+\delta p_4)(\mu-\eta)}{\mu(R_E-P_E)-\eta(T_E-S_E)},\\
\gamma &= \frac{(1-\delta p_1-p_0+\delta p_0)(\mu P_E-\eta S_E)+(p_0-\delta p_0+\delta p_4)(\mu R_E-\eta T_E)}{\mu(R_E-P_E)-\eta(T_E-S_E)},\\
p_2 &= \frac{p_1(\mu(T_E-P_E)-\eta(R_E-S_E))-(\frac{1}{\delta}+p_4)(T_E-R_E)}{\mu(R_E-P_E)-\eta(T_E-S_E)},\\
p_3 &= \frac{(\frac{1}{\delta}-p_1)(P_E-S_E)+p_4(\mu(R_E-S_E)-\eta(T_E-P_E))}{\mu(R_E-P_E)-\eta(T_E-S_E)}
\end{aligned}
\tag{30}
$$

are obtained. By substituting β and γ into Eq. (28), player Y's payoff is fixed at

$$
s_Y = \frac{(1-\delta p_1-p_0+\delta p_0)(\mu P_E-\eta S_E)+(p_0-\delta p_0+\delta p_4)(\mu R_E-\eta T_E)}{(1-\delta p_1+\delta p_4)(\mu-\eta)},
\tag{31}
$$

which is independent of the opponent's strategies $\boldsymbol{q}$. Equations (30) and (31) correspond to Eq. (10) in [52] when $\delta = 1$.

It should be noted that Equalizer does not require any condition on p_0. However, Eq. (31) indicates that the payoff that Equalizer enforces on the opponent's payoff s_Y depends on the value of p_0. Because Eq. (31) is a weighted average of $(\mu P_E - \eta S_E)/(\mu-\eta)$ and $(\mu R_E - \eta T_E)/(\mu-\eta)$ with non-negative weights, Equalizer can impose any payoff value s_Y such that $(\mu P_E - \eta S_E)/(\mu-\eta) \leq s_Y \leq (\mu R_E - \eta T_E)/(\mu-\eta)$ as long as error rate $\eta = \epsilon + \xi$ is not so high. If $(\mu P_E - \eta S_E)/(\mu-\eta)$ is enforced, it holds that $p_0 - \delta p_0 + \delta p_4 = 0$, and hence $p_4 = p_0 = 0$. If $(\mu R_E - \eta T_E)/(\mu-\eta)$ is enforced, it holds that $1 - \delta p_1 - p_0 + \delta p_0 = 0$, and hence $p_1 = p_0 = 1$.

Equalizer must satisfy the condition $0 \leq p_i \leq 1$ in Eq. (29). The existence of Equalizer strategies also depends on δ, ϵ, and ξ. We numerically find the minimum discount factor δ_c and the condition of (ϵ, ξ) for which Equalizer exists. $\delta \geq \delta_c$ is the condition for δ under which Equalizer strategies exist. Figure 1A shows δ_c when $\epsilon+\xi$ is given. We set $(T, R, P, S) = (1.5, 1, 0, -0.5)$ and excluded the case $\epsilon+\xi > 1/3$ because $T_E > R_E > P_E > S_E$ is not satisfied under the situation. Note that the effects of ϵ and ξ are the same because $\eta = \epsilon+\xi$ and $\mu = 1-\epsilon-\xi$ in Eq. (29) includes both ϵ and ξ. When there was no error ($\epsilon+\xi = 0$), δ_c was about 0.33. When the errors were $\epsilon + \xi = 0.1$ and 0.2, δ_c were about 0.52 and 0.93. As a result, we found that $\delta \geq \delta_c$ for Equalizer becomes larger as the error is increased. Figure 1B shows the possible payoff range that Equalizer can enforce to the opponent (s_Y) when corresponding $\epsilon + \xi$ is given. If there is no error ($\epsilon + \xi = 0$), Equalizer can enforce all possible payoffs between $P_E = 0$ and $R_E = 1$ as shown in Ichinose and Masuda [55]. However, as error rates become larger, the possible range of payoffs that Equalizer can enforce becomes smaller. When the errors were $\epsilon + \xi = 0.1$ and 0.2, the possible payoff ranges were $0.21 \leq s_Y \leq 0.79$ and $0.47 \leq s_Y \leq 0.53$, respectively.

Minimum correlation factor for ZD strategies with $1 \leq \chi < \infty$

We numerically calculated the minimum correlation factor χ_c for subsets of ZD strategies with $1 \leq \chi < \infty$ to exist (Fig. 2). Each curve corresponds to each δ as shown in the legend. The area surrounded by each curve and the vertical axis is the region of χ which can be utilized by the ZD strategies when $\epsilon+\xi$ is fixed. As the error $\epsilon+\xi$ becomes larger and the discount factor δ becomes smaller, the minimum correlation factor χ_c becomes larger.

III.4. Numerical examples of representative ZD and unconditional strategies under errors in repeated games

We numerically demonstrate that ZD and unconditional strategies can impose a linear relationship between the two players' payoffs while others cannot in the RPD game under errors. We take up contingent Extortion and Equalizer as the representative of ZD strategies, ALLD (always defect) as the representative of unconditional strategies, and Win-Stay-Lose-Shift (WSLS) as neither ZD nor unconditional strategies in general.

Figure 3 shows the relationship between the two players' expected payoffs per game with payoff vector $(T, R, P, S) = (1.5, 1, 0, -0.5)$. The gray quadrangle in each panel represents the feasible set of payoffs. We fixed one particular strategy for player X (vertical line) and randomly generated 1,000 strategies that satisfy $0 \leq q_0, q_1, q_2, q_3, q_4 \leq 1$ for player Y (horizontal axis). Thus, each black dot represents the payoff relationship between two players. In addition, the blue and red are the particular cases for player Y. Red is the case that player Y is ALLD and blue is the case that player Y is ALLC. We set $\delta = 1$ for Figs. 3A–D and $\delta = 0.9$ for Figs. 3E–H. In each figure, we used three error rates $\epsilon + \xi = 0, 0.1$, and 0.2.

Figures 3A and E show the case with a WSLS strategy

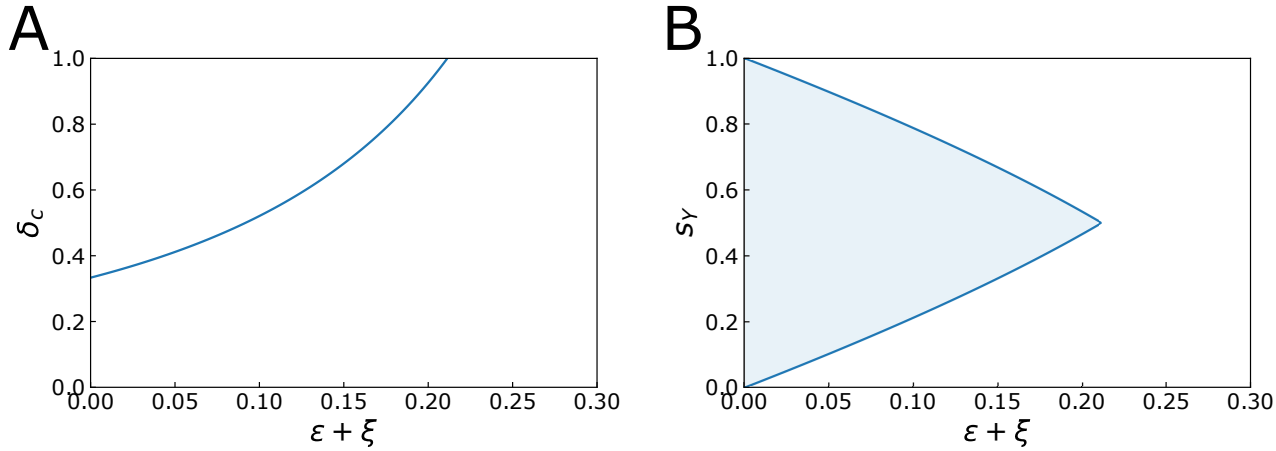

FIG. 1. (A) Minimum discount factor δ_c for Equalizer. (B) The possible payoff range that Equalizer can enforce to the opponent (s_Y) when corresponding $\epsilon + \xi$ is given.

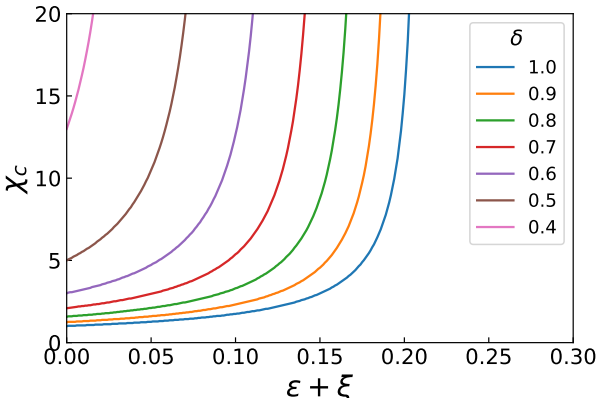

FIG. 2. Minimum correlation factor χ_c for subsets of ZD strategies with $1 \leq \chi < \infty$. $(T, R, P, S) = (1.5, 1, 0, -0.5)$. We adopted p_0 and κ so that χ_c was minimized.

vs. $1000 + 2$ strategies. In that case, $\xi = 0$ is fixed and ϵ is caried to 0, 0.1, 0.2. As WSLS strategies are neither ZD nor unconditional strategies in general, the payoff relationships are not linear irrespective of errors and a discount factor.

Figures 3B and F show the case with an contingent Extortion vs. $1000 + 2$ strategies. If there are no errors, Extortion is unbeatable against any opponent as shown by black dots. For instance, when $\delta = 1$ and $\epsilon + \xi = 0$, Extortioner $\boldsymbol{p} = (0.86, 0.77, 0.09, 0)$ which passes over (P_E, P_E) can impose a linear payoff relationship to the opponent, with the slope $\chi = 15$ [black dots in Fig. 3B]. Note that $\delta = 1$ corresponds to a no discounting game. In that case, the expected payoff of a game does not depend on p_0 because all terms which contain p_0 and q_0 in Eq. (18) vanish. In other words, the value of p_0 is arbitrary. Even if $\delta = 0.9$ and $\epsilon + \xi = 0$, Extortioner $\boldsymbol{p} = (0.955556, 0.855556, 0.1, 0; 0)$ which passes over (P_E, P_E) can impose a linear payoff relationship to the opponent, with the slope $\chi = 15$ [black dots in Fig. 3F].

However, as proved in Sec. III.2, when there are errors, only contingent Extortion can exist. Thus, there exists the region that the expected payoff of the contingent Extortion is lower than the opponent's payoff near (P_E, P_E) [see yellow-green and cyan dots in Figs. 3B and F] even though the increase of the strategies is still larger than the opponent's due to $\chi > 1$ when the opponent tries to increase his payoff.

When $\delta = 1$ and $\epsilon + \xi = 0.1$, contingent Extortioner $\boldsymbol{p} = (0.926875, 0.818125, 0.111875, 003125)$ which passes over $(P_E + 0.1, P_E + 0.1)$ can impose a linear payoff relationship to the opponent, with the slope $\chi = 15$ [yellow-green dots in Fig. 3B]. Even if $\delta = 0.9$ and $\epsilon + \xi = 0.1$, contingent Extortioner $\boldsymbol{p} = (0.941667, 0.7, 0.241667, 0; 0)$ which has the same slope $\chi = 15$ can still exist [yellow-green dots in Fig. 3F]. Nevertheless, these two contingent Extortioners' expected payoffs are lower than the opponents near (P_E, P_E). When $\delta = 1$ and $\epsilon + \xi = 0.2$, contingent Extortioner $\boldsymbol{p} = (1, 0.86, 0.14, 0)$ which passes over $(P_E + 0.2, P_E + 0.2)$ can impose a linear payoff relationship to the opponent, with the slope $\chi = 15$ [cyan dots in Fig. 3B]. However, this contingent Extortioner's payoff is lower than the opponent's near (P_E, P_E), too. When $\delta = 0.9$ and $\epsilon + \xi = 0.2$, there is no contingent Extortioner as shown in Fig. 2.

Figures 3C and G show the case with an Equalizer strategy vs. $1000 + 2$ strategies. When $\delta = 1$ and $\epsilon + \xi = 0, 0.1$, and 0.2, Equalizers $\boldsymbol{p} = (2/3, 1/3, 2/3, 1/3)$, $\boldsymbol{p} = (0.8, 0.365217, 0.634783, 0.2)$, and $\boldsymbol{p} = (0.99, 0.74, 0.26, 0.01)$ can fix the opponents' (player Y) expected payoffs at $s_Y = 0.5$ irrespective of Y's strategies, respectively [black, yellow-green, and cyan dots in Fig. 3C]. Also, when $\delta = 0.9$ and $\epsilon + \xi = 0$ and 0.1, Equalizer $\boldsymbol{p} = (2/3, 0.277778, 0.722222, 1/3; 1/2)$, and $\boldsymbol{p} = (0.833333, 0.350242, 0.649758, 1/6; 1/2)$ can fix the

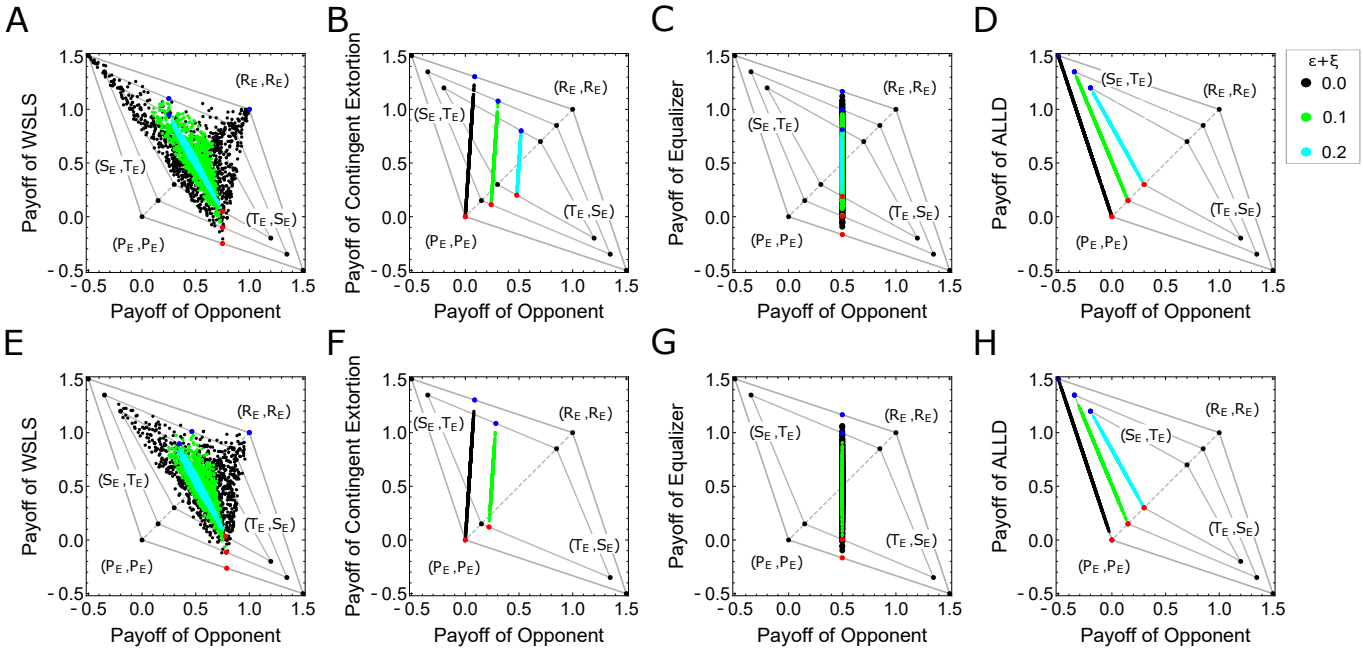

FIG. 3. The payoff relationships between two players in the RPD game under observation errors. Payoff vector: $(T, R, P, S) = (1.5, 1, 0, -0.5)$. (A,E) WSLS strategy vs. $1000 + 2$ strategies. (B,F) Contingent Extortioner strategy vs. $1000 + 2$ strategies. (C,G) Equalizer strategy vs $1000 + 2$ strategies. (D,H) ALLD strategy vs. $1000 + 2$ strategies. (A)–(D) are the case of $\delta = 1$ (no discount factor), and (E)–(H) correspond to (A)–(D) when $\delta = 0.9$, respectively.

opponents' (player Y) expected payoffs at $s_Y = 0.5$ irrespective of Y's strategies, respectively [black and yellow-green dots in Fig. 3G]. When $\delta = 0.9$ and $\epsilon + \xi = 0.2$, there exists no Equalizer as shown in Fig. 1.

Lastly, we show the case of ALLD [Figs. 3D and H]. ALLD strategy is one of the unconditional strategies where we set $r = 0$ in $\boldsymbol{p} = (r, r, r, r; r), 0 \leq r \leq 1$. As shown in Eq. (C15) and Figs. 3D and H, δ does not affect the expected payoff between both players. When $\epsilon + \xi = 0, 0.1$, and 0.2, those linear equations are $s_X + 3s_Y = 0$ (black), $s_X + 2.4s_Y - 0.51 = 0$ (yellow-green), and $s_X + 1.8s_Y - 0.84 = 0$ (cyan), respectively [53].

IV. DISCUSSION

We considered both a discount factor and observation errors in the RPD game and analytically studied the strategies that enforce linear payoff relationships in the game. First, we successfully derived the determinant form of the two players' expected payoffs even though a discount factor and observation errors are incorporated. Then, we searched for all possible strategies that enforce linear payoff relationships in the RPD game. As a result, we found that both ZD strategies and unconditional strategies are the only strategy sets to enforce the relationship to the opponent. Then, we proved that Extortion and Generous strategies no longer exist when there are errors. Finally, we numerically showed minimum discount factors for Equalizer ($\chi \to \infty$) and minimum correlation factors for other subsets of ZD strategies ($1 \leq \chi < \infty$) above which those ZD strategies exist.

We showed that ZD strategies can still exist even when discount factor δ deviates from 1. In real life, we remember interacting with other people but the interaction will conclude after a certain time. People sometimes change who they interact with because they move to other places. Young animals spend most of their time with their parents. After animals grow up and become adults, they interact with new companions. There are interactions in a short time such as mating. Our results demonstrate that strategies which unilaterally control the opponents' payoffs exist even in those limited repeated interactions.

We also showed that ZD strategies can still exist even when there are observation errors to some extent. Such noise often happens in real interactions between individuals. Thus, if ZD strategies were to exist only without noise, the applicability of ZD strategies in real problems would be quite limited. Our results are important because we expanded the applicability of ZD strategies to those problems.

Our results are limited to the two-player RPD games. Other studies have focused on n-player games [19, 26, 27, 56]. It is worth investigating games including observation errors and a discount factor for n-player games. On the other hand, regarding memory, our study only

used memory-1 strategies. A recent study revealed the role of longer memories for the evolution of cooperation, which is another direction to investigate [51].

When spatial structures are included, the different role of Extortioner has been known [16–18, 21]. Extortioners are neutral with respect to ALLDs if there are no errors. Thus, Extortioners can neutrally invade the sea of ALLDs in a spatial structure. On the other hand, the best response to Extortioners is ALLC. Once ALLC happens, the clusters of ALLC are better than those of the Extortioner. Then, cooperation is promoted. In this way, it has been demonstrated that Extortion acts as a catalyst for cooperation. Another interest is how observation errors and a discount factor affect the evolution of cooperation in a spatial setting.

We considered a model of private monitoring in direct reciprocity where signals are personally interpreted. Signals would play an important role in indirect reciprocity rather than direct reciprocity because signals are shared by many players as a social norm which affects their behaviors. As far as we know, no one has found ZD strategies in indirect reciprocity. We do not even know that we can use similar techniques for the problem. Thus, this is a potentially challenging and exciting direction of future research.

Game theory has been applied to practical problems. For instance, it has been used in the problem of selfish routing or traffic congestion. In those problems, imposing a tax on selfish terminals or drivers is one possible strategy to improve the efficiency of the total flow. As we showed, ZD strategies can induce cooperative behavior to the other players. Thus, incorporating ZD strategies instead of imposing taxes in those problems may be an effective way to improve efficiency. Our study contributes to new research directions of ZD strategies to various fields.

Appendix A: Proof of $v_1 + v_2 + v_3 + v_4 = 1$

We show that the sum of elements in the mean distribution $\boldsymbol{v} = (v_1, v_2, v_3, v_4)$ is equal to 1. We define

$$\boldsymbol{v} = (1-\delta)\boldsymbol{v}(0)\sum_{t=0}^{\infty}(\delta M)^t. \tag{A1}$$

This is another form of Eq (13). Because the sum of every row in the transition matrix M is equal to one, the sum of every row of $\sum_{t=0}^{\infty}(\delta M)^t$ is equal to $1/(1-\delta)$. The sum of vector elements in $\boldsymbol{v}(0)$ is unchanged from 1 even if the vector is multiplied by $(1-\delta)\sum_{t=0}^{\infty}(\delta M)^t$. Therefore, $v_1 + v_2 + v_3 + v_4 = 1$ holds.

Appendix B: Calculation of $\boldsymbol{v}(0) = \boldsymbol{v}^T M_0$

We show that $\boldsymbol{v}(0)$ and $\boldsymbol{v}^T M_0$ are equal. $\boldsymbol{v}$ and M_0 are defined by Eqs. (13) and (14), respectively. We calculate the matrix multiplication $\boldsymbol{v}^T M_0$.:

$$\begin{aligned}
\boldsymbol{v}^T M_0 &= (v_1, v_2, v_3, v_4)\begin{pmatrix} p_0q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \end{pmatrix} \\
&= (p_0q_0(v_1+v_2+v_3+v_4), p_0(1-q_0)(v_1+v_2+v_3+v_4), \\
&\qquad (1-p_0)q_0(v_1+v_2+v_3+v_4), (1-p_0)(1-q_0)(v_1+v_2+v_3+v_4)) \\
&= (p_0q_0, p_0(1-q_0), (1-p_0)q_0, (1-p_0)(1-q_0)) = \boldsymbol{v}(0)
\end{aligned} \tag{B1}$$

Therefore, the following holds:

$$\boldsymbol{v}(0) = \boldsymbol{v}^T M_0. \tag{B2}$$

Appendix C: Strategies that enforce $D(\boldsymbol{p},\boldsymbol{q},\alpha\boldsymbol{S_X}+\beta\boldsymbol{S_Y}+\gamma\mathbf{1})=0$.

To search for all possible strategies that make $D(\boldsymbol{p},\boldsymbol{q},\alpha\boldsymbol{S}_X+\beta\boldsymbol{S}_Y+\gamma\mathbf{1})=0$, we express Eq. (24) in component form:

$$\begin{aligned} s\begin{pmatrix} \delta(\tau p_1q_1+\epsilon p_1q_2+\epsilon p_2q_1+\xi p_2q_2)-1+p_0q_0(1-\delta)\\ \delta(\epsilon p_1q_3+\xi p_1q_4+\tau p_2q_3+\epsilon p_2q_4)+p_0q_0(1-\delta)\\ \delta(\epsilon p_3q_1+\tau p_3q_2+\xi p_4q_1+\epsilon p_4q_2)+p_0q_0(1-\delta)\\ \delta(\xi p_3q_3+\epsilon p_3q_4+\epsilon p_4q_3+\tau p_4q_4)+p_0q_0(1-\delta)\end{pmatrix}+t\begin{pmatrix}\delta(\mu p_1+\eta p_2)-1+p_0(1-\delta)\\ \delta(\eta p_1+\mu p_2)-1+p_0(1-\delta)\\ \delta(\mu p_3+\eta p_4)+p_0(1-\delta)\\ \delta(\eta p_3+\mu p_4)+p_0(1-\delta)\end{pmatrix}\\ +u\begin{pmatrix}\delta(\mu q_1+\eta q_2)-1+q_0(1-\delta)\\ \delta(\mu q_3+\eta q_4)+q_0(1-\delta)\\ \delta(\eta q_1+\mu q_2)-1+q_0(1-\delta)\\ \delta(\eta q_3+\mu q_4)+q_0(1-\delta)\end{pmatrix}+v\begin{pmatrix}\alpha R_E+\beta R_E+\gamma\\ \alpha S_E+\beta T_E+\gamma\\ \alpha T_E+\beta S_E+\gamma\\ \alpha P_E+\beta P_E+\gamma\end{pmatrix}=\mathbf{0}\end{aligned} \tag{C1}$$

By taking out $\boldsymbol{q}$ from Eq. (C1), we obtain

$$\begin{aligned}&\begin{pmatrix}\delta((s(\tau p_1+\epsilon p_2)+u\mu)q_1+(s(\epsilon p_1+\xi p_2)+u\eta)q_2)+(sp_0+u)(1-\delta)q_0\\ \delta((s(\epsilon p_1+\tau p_2)+u\mu)q_3+(s(\xi p_1+\epsilon p_2)+u\eta)q_4)+(sp_0+u)(1-\delta)q_0\\ \delta((s(\epsilon p_3+\xi p_4)+u\eta)q_1+(s(\tau p_3+\epsilon p_4)+u\mu)q_2)+(sp_0+u)(1-\delta)q_0\\ \delta((s(\xi p_3+\epsilon p_4)+u\eta)q_3+(s(\epsilon p_3+\tau p_4)+u\mu)q_4)+(sp_0+u)(1-\delta)q_0\end{pmatrix}\\ +t&\begin{pmatrix}\delta(\mu p_1+\eta p_2)-1+p_0(1-\delta)\\ \delta(\eta p_1+\mu p_2)-1+p_0(1-\delta)\\ \delta(\mu p_3+\eta p_4)+p_0(1-\delta)\\ \delta(\eta p_3+\mu p_4)+p_0(1-\delta)\end{pmatrix}+\begin{pmatrix}-s-u\\ 0\\ -u\\ 0\end{pmatrix}+v\begin{pmatrix}\alpha R_E+\beta R_E+\gamma\\ \alpha S_E+\beta T_E+\gamma\\ \alpha T_E+\beta S_E+\gamma\\ \alpha P_E+\beta P_E+\gamma\end{pmatrix}=\mathbf{0}.\end{aligned} \tag{C2}$$

Here, we search for strategies which satisfy $D(\boldsymbol{p},\boldsymbol{q},\alpha\boldsymbol{S}_X+\beta\boldsymbol{S}_Y+\gamma\mathbf{1})=0$ irrespective of Y's strategy $\boldsymbol{q}$, meaning that Eq. (C2) must hold true irrespective of $\boldsymbol{q}$. Therefore, the coefficients of each element $\boldsymbol{q}$ in Eq. (C2) must equal zero, that is, the following conditions are necessary:

$$\begin{cases}\delta(s(\epsilon p_1+\xi p_2)+u\eta) &=0,\\ \delta(s(\epsilon p_3+\xi p_4)+u\eta) &=0,\\ \delta(s(\tau p_1+\epsilon p_2)+u\mu) &=0,\\ \delta(s(\tau p_3+\epsilon p_4)+u\mu) &=0,\\ \delta(s(\epsilon p_1+\tau p_2)+u\mu) &=0,\\ \delta(s(\xi p_1+\epsilon p_2)+u\eta) &=0,\\ \delta(s(\xi p_3+\epsilon p_4)+u\eta) &=0,\\ \delta(s(\epsilon p_3+\tau p_4)+u\mu) &=0,\\ (sp_0+u)(1-\delta) &=0.\end{cases} \tag{C3}$$

When Eq. (C3) holds, the first terms of Eq. (C2) are eliminated and we obtain

$$t\begin{pmatrix}\delta(\mu p_1+\eta p_2)-1+p_0(1-\delta)\\ \delta(\eta p_1+\mu p_2)-1+p_0(1-\delta)\\ \delta(\mu p_3+\eta p_4)+p_0(1-\delta)\\ \delta(\eta p_3+\mu p_4)+p_0(1-\delta)\end{pmatrix}+\begin{pmatrix}-s-u\\ 0\\ -u\\ 0\end{pmatrix}+v\begin{pmatrix}\alpha R_E+\beta R_E+\gamma\\ \alpha S_E+\beta T_E+\gamma\\ \alpha T_E+\beta S_E+\gamma\\ \alpha P_E+\beta P_E+\gamma\end{pmatrix}=\mathbf{0}. \tag{C4}$$

If there exist real numbers, s,t,u,v,α,β, and γ such that Eqs. (C3) and (C4) are satisfied simultaneously, $D(\boldsymbol{p},\boldsymbol{q},\alpha\boldsymbol{S}_X+\beta\boldsymbol{S}_Y+\gamma\mathbf{1})=0$ holds irrespective of $\boldsymbol{q}$. We first solve Eq. (C3). After some calculations, Eq. (C3)

becomes

$$\begin{cases} \delta s(\epsilon-\xi)(p_1-p_2) & =0, \\ \delta s(\epsilon-\xi)(p_3-p_4) & =0, \\ \delta s(1-3\epsilon-\xi)(p_1-p_2) & =0, \\ \delta s(1-3\epsilon-\xi)(p_3-p_4) & =0, \\ \delta(s(\epsilon p_1+\tau p_2)+u\mu) & =0, \\ \delta(s(\xi p_1+\epsilon p_2)+u\eta) & =0, \\ \delta(s(\xi p_3+\epsilon p_4)+u\eta) & =0, \\ \delta(s(\epsilon p_3+\tau p_4)+u\mu) & =0, \\ (sp_0+u)(1-\delta) & =0. \end{cases} \tag{C5}$$

When we solve the first four equations, we obtain (1) $s=0$, (2) $\epsilon-\xi=0$ and $1-3\epsilon-\xi=0$, (3) $p_1-p_2=0$ and $p_3-p_4=0$. We further analyze whether these solutions satisfy the last four equations and Eq. (C4) by dividing them into four cases as follows.

Case (1): $s=0$

In this case, we substitute $s=0$ into Eq. (C5) to obtain

$$\begin{cases} u\mu & =0, \\ u\eta & =0, \\ u(1-\delta) & =0. \end{cases} \tag{C6}$$

The equations $\mu=0$ and $\eta=0$ do not hold at the same time due to $\mu=1-\epsilon-\xi$ and $\eta=\epsilon+\xi$. Therefore, one of the solutions of Eq. (C5) is $s=0$ and $u=0$. Next, we check whether this solution satisfies Eq. (C4). We substitute $s=0$ and $u=0$ into Eq. (C4) to obtain

$$t\begin{pmatrix} \delta(\mu p_1+\eta p_2)-1+p_0(1-\delta) \\ \delta(\eta p_1+\mu p_2)-1+p_0(1-\delta) \\ \delta(\mu p_3+\eta p_4)+p_0(1-\delta) \\ \delta(\eta p_3+\mu p_4)+p_0(1-\delta) \end{pmatrix}+v\begin{pmatrix} \alpha R_E+\beta R_E+\gamma \\ \alpha S_E+\beta T_E+\gamma \\ \alpha T_E+\beta S_E+\gamma \\ \alpha P_E+\beta P_E+\gamma \end{pmatrix}=\mathbf{0}. \tag{C7}$$

Here, when we set $t=0$, either equation

$$v=0 \tag{C8}$$

or

$$\begin{pmatrix} \alpha R_E+\beta R_E+\gamma \\ \alpha S_E+\beta T_E+\gamma \\ \alpha T_E+\beta S_E+\gamma \\ \alpha P_E+\beta P_E+\gamma \end{pmatrix}=\mathbf{0} \tag{C9}$$

must hold. When we set $v=0$, we obtain the trivial solution $(s,t,u,v)=(0,0,0,0)$. Also, when we solve Eq. (C9), we obtain the trivial solution $(\alpha,\beta,\gamma)=(0,0,0)$. Hence, we do not have to consider the case of $t=0$. Therefore, in the following, we only consider $t\neq 0$. Replacing constants $-\alpha v/t$, $-\beta v/t$, and $-\gamma v/t$ with α, β, and γ, we obtain,

$$\begin{aligned} \delta(\mu p_1+\eta p_2)-1+p_0(1-\delta) &= \alpha R_E+\beta R_E+\gamma, \\ \delta(\eta p_1+\mu p_2)-1+p_0(1-\delta) &= \alpha S_E+\beta T_E+\gamma, \\ \delta(\mu p_3+\eta p_4)+p_0(1-\delta) &= \alpha T_E+\beta S_E+\gamma, \\ \delta(\eta p_3+\mu p_4)+p_0(1-\delta) &= \alpha P_E+\beta P_E+\gamma. \end{aligned} \tag{C10}$$

If there exist α,β, and γ satisfying Eq. (C10), there must be solutions that Eq. (24) holds. This solution corresponds to ZD strategies with observation errors and a discount factor. This is consistent with Eq. (6) in [52] when $\delta=1$.

Case (2): $\epsilon - \xi = 0$ **and** $1 - 3\epsilon - \xi = 0$

In this case, the equations $\epsilon - \xi = 0$ and $1 - 3\epsilon - \xi = 0$ lead to $\epsilon = 1/4$ and $\xi = 1/4$. When $\epsilon = 1/4$ and $\xi = 1/4$, the expected payoffs $R_E = 1/2(R + S), S_E = 1/2(R + S), T_E = 1/2(T + P)$, and $P_E = 1/2(T + P)$ hold, which do not satisfy the condition of the prisoner's dilemma game: $T_E > R_E > P_E > S_E$. Hence, we can exclude this solution.

Case (3): $p_1 - p_2 = 0$ **and** $p_3 - p_4 = 0$

In this case, we substitute $p_1 - p_2 = 0$ and $p_3 - p_4 = 0$ into Eq. (C5) to obtain

$$\begin{cases} \delta\mu(sp_1 + u) & = 0, \\ \delta\eta(sp_1 + u) & = 0, \\ \delta\eta(sp_3 + u) & = 0, \\ \delta\mu(sp_3 + u) & = 0, \\ (sp_0 + u)(1 - \delta) & = 0. \end{cases} \tag{C11}$$

Because the equations $\mu = 0$ and $\eta = 0$ do not hold at the same time and $\delta \neq 0$, we obtain

$$\begin{cases} sp_1 + u & = 0, \\ sp_3 + u & = 0, \\ (sp_0 + u)(1 - \delta) & = 0. \end{cases} \tag{C12}$$

Therefore, we obtain two solutions $p_0 = p_1 = p_2 = p_3 = p_4 = -u/s$ or $p_1 = p_2 = p_3 = p_4 = -u/s$ and $\delta = 1$. Both solutions are called unconditional strategies [14, 55]. The former represents unconditional strategies in the case of $\delta \neq 1$. The latter represents unconditional strategies in the case of $\delta = 1$. Next, we check whether this solution satisfies Eq. (C4). We substitute the former solution $p_0 = p_1 = p_2 = p_3 = p_4$ and $u = -sp_0$ into Eq. (C4) to obtain

$$s \begin{pmatrix} p_0 - 1 \\ 0 \\ p_0 \\ 0 \end{pmatrix} + t \begin{pmatrix} \delta(\mu p_0 + \eta p_0) - 1 + p_0(1 - \delta) \\ \delta(\eta p_0 + \mu p_0) - 1 + p_0(1 - \delta) \\ \delta(\mu p_0 + \eta p_0) + p_0(1 - \delta) \\ \delta(\eta p_0 + \mu p_0) + p_0(1 - \delta) \end{pmatrix} + v \begin{pmatrix} \alpha R_E + \beta R_E + \gamma \\ \alpha S_E + \beta T_E + \gamma \\ \alpha T_E + \beta S_E + \gamma \\ \alpha P_E + \beta P_E + \gamma \end{pmatrix} = \mathbf{0}. \tag{C13}$$

According to $\mu + \eta = 1$, we obtain

$$s \begin{pmatrix} p_0 - 1 \\ 0 \\ p_0 \\ 0 \end{pmatrix} + t \begin{pmatrix} p_0 - 1 \\ p_0 - 1 \\ p_0 \\ p_0 \end{pmatrix} + v \begin{pmatrix} \alpha R_E + \beta R_E + \gamma \\ \alpha S_E + \beta T_E + \gamma \\ \alpha T_E + \beta S_E + \gamma \\ \alpha P_E + \beta P_E + \gamma \end{pmatrix} = \mathbf{0}. \tag{C14}$$

There exist real numbers $s, t, u, v, \alpha, \beta$, and γ which satisfy Eq. (C14) as follows:

$$\begin{aligned} s &= \frac{v\alpha(S_E(-P_E - R_E + S_E) + T_E(P_E + R_E - T_E))}{(1 - p_0)(P_E - S_E) + p_0(T_E - R_E)}, \\ t &= \frac{v\alpha(S_E(2P_E - S_E + p_0[-P_E - R_E + S_E]) + T_E(-2P_E + T_E + p_0[P_E + R_E - T_E]))}{(1 - p_0)(P_E - S_E) + p_0(T_E - R_E)}, \\ u &= -sp_0, \\ \beta &= \frac{\alpha((1 - p_0)(T_E - P_E) + p_0(R_E - S_E))}{(1 - p_0)(P_E - S_E) + p_0(T_E - R_E)}, \\ \gamma &= \frac{\alpha(S_E - T_E)((-1 + p_0)^2 P_E + p_0(1 - p_0)(T_E + S_E) + p_0^2 R_E)}{(1 - p_0)(P_E - S_E) + p_0(T_E - R_E)}, \\ &\forall v, \alpha. \end{aligned} \tag{C15}$$

Finally, we substitute the latter solution $p_1 = p_2 = p_3 = p_4 = -u/s$ and $\delta = 1$ into Eq (C4), we obtain the same real numbers $s, t, u, v, \alpha, \beta$, and γ in the case of the former solution.

This strategy set corresponds to unconditional strategies $\boldsymbol{p} = (r, r, r, r; r), 0 \leq r \leq 1$. Therefore, the unconditional strategies enforce a linear payoff relationship in the RPD game with both observation errors and a discount factor because there exist real numbers $s, t, u, v, \alpha, \beta$, and γ such that Eqs. (C4) and (C5) are satisfied.

Appendix D: Calculation of Sec. III.2

We show that neither Extortion nor Generous strategies exist when there are errors. From Eq. (27), we can rewrite ZD strategies as follows:

$$\begin{aligned}
&\delta(\mu p_1 + \eta p_2) + p_0(1-\delta) = 1 - \phi(\chi-1)(R_E - \kappa),\\
&\delta(\eta p_1 + \mu p_2) + p_0(1-\delta) = 1 - \phi[(\chi-1)(R_E-\kappa) + (R_E - S_E) + \chi(T_E - R_E)],\\
&\delta(\mu p_3 + \eta p_4) + p_0(1-\delta) = \phi[(\chi-1)(\kappa - P_E) + (T_E - P_E) + \chi(P_E - S_E)],\\
&\delta(\eta p_3 + \mu p_4) + p_0(1-\delta) = \phi(\chi-1)(\kappa - P_E).
\end{aligned} \tag{D1}$$

We assume $\chi \geq 1$ because Extortion and Generous are considered. We also assume $\eta > 0$ because we consider errors.

If we solve those equations for p_1, p_2, p_3 and p_4, respectively, we obtain

$$\begin{aligned}
\delta p_1 + (1-\delta)p_0 &= 1 - \phi\left((\chi-1)(R_E-\kappa) - \frac{\eta}{\mu-\eta}((R_E-S_E) + \chi(T_E-R_E))\right),\\
\delta p_2 + (1-\delta)p_0 &= 1 - \phi\left((\chi-1)(R_E-\kappa) + \frac{\mu}{\mu-\eta}((R_E-S_E) + \chi(T_E-R_E))\right),\\
\delta p_3 + (1-\delta)p_0 &= \phi\left((\chi-1)(\kappa-P_E) + \frac{\mu}{\mu-\eta}((T_E-P_E) + \chi(P_E-S_E))\right),\\
\delta p_4 + (1-\delta)p_0 &= \phi\left((\chi-1)(\kappa-P_E) - \frac{\eta}{\mu-\eta}((T_E-P_E) + \chi(P_E-S_E))\right).
\end{aligned} \tag{D2}$$

In those equations, when $\delta < 1$ and $\phi = 0$ are given, p_1, p_2, p_3 and p_4 do not exist. When $\delta = 1$, $\phi = 0$ is formally allowed, but produces only the singular strategy (1,1,0,0). Thus, we consider the case of $\phi \neq 0$.

We substitute $\kappa = P_E$ into the third and fourth equations in Eq. (D2) to obtain

$$\begin{aligned}
\delta p_3 + (1-\delta)p_0 &= \phi\left(\frac{\mu}{\mu-\eta}((T_E-P_E) + \chi(P_E-S_E))\right),\\
\delta p_4 + (1-\delta)p_0 &= \phi\left(-\frac{\eta}{\mu-\eta}((T_E-P_E) + \chi(P_E-S_E))\right).
\end{aligned} \tag{D3}$$

For those two equations, the left sides must be greater than or equal to zero. Thus, the right sides need to be positive due to $\phi \neq 0$. When $\mu > \eta$, we obtain $\phi > 0$ from the first equation while $\phi < 0$ from the second equation. Thus, there is no ϕ which satisfies the condition in this case. When $\mu < \eta$, we obtain $\phi < 0$ from the first equation while $\phi > 0$ from the second equation. Thus, there is also no ϕ which satisfies the condition in this case. We exclude the case of $\mu = \eta$ because this yields $R_E = S_E$ and $T_E = P_E$ which is not the RPD. Therefore, there is no ZD with $\kappa = P_E$, which corresponds to Extortion, when there are errors.

We substitute $\kappa = R_E$ into the first and second equations in Eq. (D2) to obtain

$$\begin{aligned}
\delta p_1 + (1-\delta)p_0 &= 1 - \phi\left(-\frac{\eta}{\mu-\eta}((R_E-S_E) + \chi(T_E-R_E))\right)\\
\delta p_2 + (1-\delta)p_0 &= 1 - \phi\left(\frac{\mu}{\mu-\eta}((R_E-S_E) + \chi(T_E-R_E))\right).
\end{aligned} \tag{D4}$$

In addition, from $0 \leq p_i \leq 1$, we obtain

$$\begin{aligned}
(1-\delta)(1-p_0) &\leq \phi\left(-\frac{\eta}{\mu-\eta}((R_E-S_E) + \chi(T_E-R_E))\right) \leq (1-\delta)(1-p_0) + \delta\\
(1-\delta)(1-p_0) &\leq \phi\left(\frac{\mu}{\mu-\eta}((R_E-S_E) + \chi(T_E-R_E))\right) \leq (1-\delta)(1-p_0) + \delta.
\end{aligned} \tag{D5}$$

For the inequalities, the middle terms must be positive due to $\phi \neq 0$. When $\mu > \eta$, we obtain $\phi < 0$ from the first equation while $\phi > 0$ from the second equation. Thus, there is no ϕ which satisfies the condition in this case. When $\mu < \eta$, we obtain $\phi > 0$ from the first equation while $\phi < 0$ from the second equation. Thus, there is also no ϕ which satisfies the condition in this case. We exclude the case of $\mu = \eta$ because this yields $R_E = S_E$ and $T_E = P_E$ which is not the RPD. Therefore, there is no ZD with $\kappa = R_E$, which corresponds to Generous, when there are errors.

ACKNOWLEDGMENTS

This study was partly supported by the HAYAO NAKAYAMA Foundation for Science & Technology and Culture and JSPS KAKENHI Grants No. JP19K04903 and No. JP19KK0262 (G.I.).

AUTHOR CONTRIBUTIONS

A.M. and G.I. designed the study and wrote the paper. A.M. conducted mathematical and numerical analyses. G.I. checked the calculation and provided the interpretation of the result.

COMPETING INTERESTS

The authors declare no competing interests.

[1] J. Mailath and L. Samuelson, *Repeated Games and Reputation* (Oxford University Press, Oxford, 2006).
[2] R. L. Trivers, Q. Rev. Biol. **46**, 35 (1971).
[3] M. A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).
[4] K. Sigmund, *The Calculus of Selfishness* (Princeton University Press, Princeton, NJ, 2010).
[5] J. M. Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, UK, 1982).
[6] M. A. Nowak and K. Sigmund, Nature (London) **355**, 250 (1992).
[7] M. Nowak and K. Sigmund, Nature (London) **364**, 56 (1993).
[8] D. Kraines and V. Kraines, Theory Decis. **35**, 107 (1993).
[9] W. H. Press and F. J. Dyson, Proc. Natl. Acad. Sci. USA **109**, 10409 (2012).
[10] A. J. Stewart and J. B. Plotkin, Proc. Natl. Acad. Sci. USA **109**, 10134 (2012).
[11] E. Akin, in Ergodic Theory, Advances in Dynamical Systems **(de Gruyter, Berlin, 2016)**, p. 77.
[12] C. Adami and A. Hintze, Nat. Comm. **4**, 2193 (2013).
[13] C. Hilbe, M. A. Nowak, and K. Sigmund, Proc. Natl. Acad. Sci. USA **110**, 6913 (2013).
[14] C. Hilbe, M. A. Nowak, and A. Traulsen, PLoS ONE **8**, e77886 (2013).
[15] J. Chen and A. Zinger, J. Theor. Biol. **357**, 46 (2014).
[16] A. Szolnoki and M. Perc, Phys. Rev. E **89**, 022804 (2014).
[17] A. Szolnoki and M. Perc, Sci. Rep. **4**, 5496 (2014).
[18] Z. X. Wu and Z. Rong, Phys. Rev. E **90**, 062102 (2014).
[19] C. Hilbe, B. Wu, A. Traulsen, and M. A. Nowak, J. Theor. Biol. **374**, 115 (2015).
[20] J. Liu, Y. Li, C. Xu, and P. M. Hui, Physica A **430**, 81 (2015).
[21] X. Xu, Z. Rong, Z. X. Wu, T. Zhou, and C. K. Tse, Phys. Rev. E **95**, 052302 (2017).
[22] J. Wang and J. Guo, Chaos, Solitons & Fractals **119**, 263 (2019).
[23] A. J. Stewart and J. B. Plotkin, Proc. Natl. Acad. Sci. USA **110**, 15348 (2013).
[24] Y. Mao, X. Xu, Z. Rong, and Z. X. Wu, EPL **122**, 50005 (2018).
[25] X. Xu, Z. Rong, Z. Tian, and Z. X. Wu, Neurocomputing **350**, 195 (2019).
[26] C. Hilbe, B. Wu, A. Traulsen, and M. A. Nowak, Proc. Natl. Acad. Sci. USA **111**, 16425 (2014).
[27] L. Pan, D. Hao, Z. Rong, and T. Zhou, Sci. Rep. **5**, 13096 (2015).
[28] M. Milinski, C. Hilbe, D. Semmann, R. Sommerfeld, and J. Marotzke, Nat. Comm. **7**, 10915 (2016).
[29] A. J. Stewart, T. L. Parsons, and J. B. Plotkin, Proc. Natl. Acad. Sci. USA **113**, E7003 (2016).
[30] A. McAvoy and C. Hauert, Proc. Natl. Acad. Sci. USA **113**, 3573 (2016).
[31] A. McAvoy and C. Hauert, Theor. Popul. Biol. **113**, 13 (2017).
[32] M. A. Taha and A. Ghoneim, Appl. Math. Comput. **369** , 124862 (2019).
[33] A. Engel and A. Feigel, Phys. Rev. E **98**, 012415 (2018).
[34] C. Hilbe, T. Röhl, and M. Milinski, Nat. Comm. **5**, 3976 (2014).
[35] Z. Wang, Y. Zhou, J. W. Lien, J. Zheng, and B. Xu, Nat. Comm. **7**, 11125 (2016).
[36] C. Hilbe, K. Hagel, and M. Milinski, PLoS ONE **11**, e0163867 (2016).
[37] L. Becks and M. Milinski, Nat. Comm. **10**, 783 (2019).
[38] E. Akin, Games **6**, 175 (2015).
[39] C. Hilbe, K. Chatterjee, and M. A. Nowak, Nat. Hum. Behav. **2**, 469 (2018).
[40] Q. Hu, S. Wang, X. Cheng, L. Ma, and R. Bie, IEEE Trans. Inf. Forensic Secur. , (2019).
[41] Q. Hu, S. Wang, P. Ma, X. Cheng, W. Lv, and R. Bie, IEEE Trans. Knowl. Data Eng. , (2019).
[42] C. Tang, X. Li, M. Cao, Z. Zhang, and X. Yu, IEEE Internet Things J. **6**, 8589 (2019).
[43] M. Kandori, J. Econ. Theor. **102**, 1 (2002).
[44] M. A. Nowak, K. Sigmund, and E. El-Sedy, J. Math. Biol. **33**, 703 (1995).
[45] D. Fudenberg, D. G. Rand, and A. Dreber, Am. Econ. Rev. **102**, 720 (2012).
[46] D. Fudenberg, D. Levine, and E. Maskin, Econometrica **62**, 997 (1994).
[47] T. Sekiguchi, J. Econ. Theor. **76**, 345 (1997).

[48] M. Barlo, G. Carmona, and H. Sabourian, J. Econ. Theor. **144**, 312 (2009).
[49] G. J. Mailath and S. Morris, J. Econ. Theor. **102**, 189 (2002).
[50] G. J. Mailath and W. Olszewski, Gam. Econ. Behav. **71**, 174 (2011).
[51] C. Hilbe, L. A. Martinez-Vaqueroand, K. Chatterjee, and M. A. Nowak, Proc. Natl. Acad. Sci. USA **114**, 4715 (2017).
[52] D. Hao, Z. Rong, and T. Zhou, Phys. Rev. E **91**, 052803 (2015).
[53] A. Mamiya and G. Ichinose, J. Theor. Biol. **477**, 63 (2019).
[54] C. Hilbe, A. Traulsen, and K. Sigmund, Games Econ. Behav. **92**, 41 (2015).
[55] G. Ichinose and N. Masuda, J. Theor. Biol. **438**, 61 (2018).
[56] A. Govaert and M. Cao, arXiv:1910.07858.