PHYSICAL REVIEW E **91**, 052803 (2015)

Extortion under uncertainty: Zero-determinant strategies in noisy games

Dong Hao,[1] Zhihai Rong,[1] and Tao Zhou[1,2,*]

[1]*CompleX Lab, Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 611731, P.R. China*

[2]*Big Data Research Center, University of Electronic Science and Technology of China, Chengdu 611731, P.R. China*

(Received 1 February 2015; published 11 May 2015)

Repeated game theory has been one of the most prevailing tools for understanding long-running relationships, which are the foundation in building human society. Recent works have revealed a new set of "zero-determinant" (ZD) strategies, which is an important advance in repeated games. A ZD strategy player can exert unilateral control on two players' payoffs. In particular, he can deterministically set the opponent's payoff or enforce an unfair linear relationship between the players' payoffs, thereby always seizing an advantageous share of payoffs. One of the limitations of the original ZD strategy, however, is that it does not capture the notion of robustness when the game is subjected to stochastic errors. In this paper, we propose a general model of ZD strategies for noisy repeated games and find that ZD strategies have high robustness against errors. We further derive the pinning strategy under noise, by which the ZD strategy player coercively sets the opponent's expected payoff to his desired level, although his payoff control ability declines with the increase of noise strength. Due to the uncertainty caused by noise, the ZD strategy player cannot ensure his payoff to be permanently higher than the opponent's, which implies *dominant extortions* do not exist even under low noise. While we show that the ZD strategy player can still establish a novel kind of extortions, named *contingent extortions*, where any increase of his own payoff always exceeds that of the opponent's by a fixed percentage, and the conditions under which the contingent extortions can be realized are more stringent as the noise becomes stronger.

DOI: 10.1103/PhysRevE.91.052803 PACS number(s): 89.75.Fb, 02.50.Le, 89.20.Ff, 87.23.−n

I. INTRODUCTION

Repeated games have been representative to explore the agents' long-run relationships, which help us in understanding how cooperation and competition might arise among agents with selfish objectives. Extensive literatures have by now utilized repeated games as a basic component to analyze economic behaviors, political science, evolutionary dynamics, as well as multiagent systems [1]. It has been commonly accepted that in such games there is no simple ultimatum strategy whereby one player can simply occupy an unfair share of the payoffs. However, Press and Dyson's discovery of "zero-determinant" (ZD) strategies illuminates a new starting point [2]. They show that in an iterated prisoner's dilemma, it is possible for a player (named the *ZD player* for short) to unilaterally enforce a linear relationship between his and the opponent's payoff, thereby deterministically setting the expected payoff of the opponent to a fixed value or ensuring that, when the opponent tries to increase his payoff, he will always increase the ZD player's payoff even more. The discovery of ZD strategies is a milestone along the way to fundamentally understanding the underlying norms of social interactions and how different strategies correlate with each other [3,4]. It provides us with a powerful but succinct framework for motivating and sustaining the cooperation required for any society, as well as for controlling the damage done by unscrupulous or mischievous agents.

ZD strategies have thus attracted considerable attention and been incorporated successfully into a wide array of research, ranging from theoretical game research to real-world experimental studies [5]. Among the subsequent research, Roemheld generalizes ZD strategies for all symmetric bimatrix games as well as for the "battle of the sexes," which is the most common example for asymmetric games [6]. Akin explores a broader space of strategies by extending Press-Dyson theorem and obtains cooperation-enforcing good strategies [7]. Thereafter, Stewart and Plotkin, as well as Hilbe *et al.*, identify the intersection of ZD strategies and good strategies, named generous ZD strategies, which not only control the payoffs but also cooperate with others and forgive defecting opponents, leading the game towards a win-win situation [8,9]. Chen and Zinger analyze the robustness of ZD strategies against evolutionary players and prove that there always exist evolutionary paths for the ZD player to obtain the maximum payoff [10]. Press and Dyson's work can be further generalized to multiplayer ZD strategies for investigating various social dilemmas, and new features and constraints related to participant number and payoff structure have been revealed and the impact of ZD alliance in multiplayer games has been studied [11,12]. Furthermore, there is also extensive work investigating the significance of ZD strategies in evolutionary game theory and in social networks [8,9,12–18]. Although initially the evolutionary instability was found for extortion strategies [14], later it is proved that the generous strategies finally dominate in the population and are stable in an evolutionary sense [8,15]. The above theoretical studies also have been implemented in real-world social experiments, and it is confirmed that extorting others has limited prospects and, in the long run, generosity is more profitable [13].

By now, how the ZD strategies perform in realistic noisy games is still an open problem. As in Stewart and Plotkin's commentary to Press and Dyson's work [3], one of the key questions is as follows: How do ZD strategies fare in iterated games in the presence of noise? Since stochastic perturbations due to observation errors, action mistakes,

*zhutou@ustc.edu

biological mutations, and other chance events are common and inevitable in reality, it is of great importance to extensively investigate the strategies and solutions in games theory at the presence of noise. However, the majority of known results on game theory [19], as well as those related to ZD strategies [5], are obtained in a perfect environment without any noise. Actually, the analyses of noisy repeated games have been long-standing challenges and are at the cutting edge of research on game theory and social interactions [19–26]. The errors in noisy repeated games usually fall into two categories [20]. The first kind is where players' actions are often observed with errors, which can be called *perception errors*: Someone who claims they worked hard or that they were too busy to give help may or may not be telling the truth; similarly, awkward results sometimes accidentally follow good behaviors [21]. The second kind is where players may wrongly take an action. This is categorized into *implementation errors* (or *action errors* in the literature): One player has an intended action but may accidentally choose another action due to interferences; this is also described by the well-known notion of "trembling hands" [22].

To explore noisy games, which is new territory for ZD strategies, we propose a general framework of ZD strategies in noisy repeated games and show the implementable for a unilateral payoff control. Since repeated games with perception errors are the most stringent case [1,19], our analysis focuses primarily on this scenario, and it can be easily extended to repeated games with implementation errors. It is found that ZD strategies present strong robustness against noise. Even in environments with perception or implementation errors, a player can still enforce a linear relationship between the two players' payoffs. Under noisy repeated games, we classify the ZD strategies into three subsets, (i) pinning strategies, (ii) contingent extortion strategies, and (iii) dominant extortion strategies. Following the pinning strategy, the ZD player can unilaterally set the opponent's payoff to his or desirable level, although the difficulty for realizing such payoff control increases as the noise becomes stronger. Furthermore, we prove that since the noise brings uncertainty and risk to the ZD player, he cannot perfectly secure his payoff to be always greater than that of the opponent. That is to say, dominant extortions do not exist even when the noise strength is low. Nevertheless, the ZD player can still extort the opponent and grab the achievements of him, in the sense that as long as the opponent tries to improve his payoff, he will improve the ZD player's payoff even more, and the opponent can only maximize his payoff by fully cooperating. At that point, both players' payoffs are maximized but the ZD player outperforms the opponent. Using such a strategy, the ZD player can extort the opponent, but he also suffers a risk of being outperformed by the opponent. Therefore we call such strategy the *contingent extortion strategy*. Our analyses of extortion in noisy games imply that errors expose the ZD player to uncertainty and risk of losing, while the mischievous manipulation and the unusual control still stubbornly persist. The results of our study can be utilized both to propose a generalized framework for the ZD strategy paradigm that has characterized much of the recent literatures and to provide a unilateral payoff control scheme for a larger class of noisy repeated games where payoff control is of great significance but has barely been studied.

II. NOISY REPEATED GAME

Consider two players engaged in an iterated prisoner's dilemma (IPD) game. In each stage, each player $i \in \{X,Y\}$ takes an action $a_i \in \{C,D\}$. Each player cannot directly see what action the opponent has taken but only observes a private signal $\omega_i \in \{g,b\}$, where g and b denote good and bad signals, respectively. Each player's signal ω_i is a stochastic variable, affected not only by the two players' actions but also by the noises (random errors) from the environment. Given the actions, every possible signal profile occurs with a positive probability $\pi(\omega|\mathbf{a})$, where $\omega = \{\omega_X,\omega_Y\}$ and $\mathbf{a} = \{a_X,a_Y\}$ are the observed signal profile and the action profile, respectively. In each stage, if player Y chooses $a_Y = C$ (or $a_Y = D$) but X observes $\omega_X = b$ (or $\omega_X = g$), it means an *error* occurs. More precisely, denote ε as the commonly known probability that an error occurs to exactly one player and denote ξ as the probability that errors occur to both players. Then the probability that neither player has an error is $1-2\varepsilon-\xi$. In particular, this setting captures both the case with independent errors and the case with correlated errors. If both players take action C, then $\pi(g,g|CC) = 1-2\varepsilon-\xi$, $\pi(g,b|CC) = \pi(b,g|CC) = \varepsilon$, and $\pi(b,b|CC) = \xi$. The following tables summarize the signal distributions under all action profiles. Based on the action and privately observed signal, for a player X, his private outcome in each stage game is a tuple $(a_X,\omega_X) \in \{Cg,Cb,Dg,Db\}$. The signal distributions for different action profiles are summarized in Table I. Note that this differs from games without noise, where both players' outcomes are identical and are just action profiles.

Since the stochastic changes of the environment, as well as the opponent's action, is jointly involved in the signals, the *realized payoff* for each player depends only on the action he chose and the signal he received, denoted as $u_i(a_i,\omega_i)$ [1,19,23]. Assume that the realized stage payoff follows the prisoner's dilemma, such that $u_i(C,g) = 1$, $u_i(C,b) = -L$, $u_i(D,g) = 1+G$, and $u_i(D,b) = 0$, where L and G are positive variables. According to the general framework in Ref. [23], in each stage, player i's *expected payoff* when two players have an action profile $\mathbf{a}$ is derived as

$$f_i(\mathbf{a}) = \sum_{\omega} u_i(a_i,\omega_i)\pi(\omega|\mathbf{a}), \quad (1)$$

TABLE I. Signal distributions for different action profiles.

CC	$\omega_Y = g$	$\omega_Y = b$
$\omega_X = g$	$1-2\varepsilon-\xi$	ε
$\omega_X = b$	ε	ξ
DC	$\omega_Y = g$	$\omega_Y = b$
$\omega_X = g$	ε	$1-2\varepsilon-\xi$
$\omega_X = b$	ξ	ε
CD	$\omega_Y = g$	$\omega_Y = b$
$\omega_X = g$	ε	ξ
$\omega_X = b$	$1-2\varepsilon-\xi$	ε
DD	$\omega_Y = g$	$\omega_Y = b$
$\omega_X = g$	ξ	ε
$\omega_X = b$	ε	$1-2\varepsilon-\xi$

such that $f_i(\mathbf{a}$ is the expected value over all possible signals, dependent on the two players' actions. When players cooperate with each other, the errors may occur to both players, and the realized stage payoff will be $-L$. However, when ε and ξ are small, the expected payoff over different possible signals may still be high. The expected payoffs under different action profiles CC, CD, DC, and DD are denoted as R_E, S_E, T_E, and P_E, which can be derived according to Eq. (1) as $R_E = 1-(L+1)(\varepsilon+\xi)$, $S_E = -L+(1+L)(\varepsilon+\xi)$, $T_E = (1+G)(1-\varepsilon-\xi)$, and $P_E = (1+G)(\varepsilon+\xi)$, respectively. Then player X's expected stage payoff vector is denoted as $\mathbf{U_X} = (R_E,S_E,T_E,P_E)$ and player Y's is denoted as $\mathbf{U_Y} = (R_E,T_E,S_E,P_E)$.

We concentrate on the memory-one strategies where each player sets his strategy only according to the single previous outcome [2,24,27]. Denote the probabilities that player X will cooperate under his previous outcomes Cg, Cb, Dg, and Db as p_1, p_2, p_3, and p_4 and the probabilities that Y will cooperate under his previous outcomes Cg, Cb, Dg, and Db are q_1, q_2, q_3, and q_4. The joint actions of the two players are the *states* of the game, and the two players' probabilistic strategies as well as the noise structure jointly determine the transition rule of the states. Note that the observation errors only change the transition probabilities but never change the real state space of the game, which is still $\{CC,CD,DC,DD\}$. For example, if the old state is CC, the probability that the state transits to a new joint state CD will be $(1-2\varepsilon-\xi)p_1(1-q_1)+\varepsilon p_1(1-q_2)+\varepsilon p_2(1-q_1)+\xi p_2(1-q_2)$, where $(1-2\varepsilon-\xi)p_1(1-q_1)$ is the probability that both players observe correct signals and player X takes action C while player Y takes action D in the new stage; $\varepsilon p_1(1-q_2)$ and $\varepsilon p_2(1-q_1)$ are the probabilities one player has an observation error and player X takes C and player Y takes D; and $\xi p_2(1-q_2)$ is the probability that both players have observation errors and player X takes C and player Y takes D. Let $\tau = 1-2\varepsilon-\xi$, the derivation of the transition probability from state CC to state CD is depicted in Fig. 1. This figure illustrates that the noise decomposes the state CC into four combinations of private outcomes, namely (Cg,Cg), (Cg,Cb), (Cb,Cg), and (Cb,Cb). Following in the same way, the state transition matrix $\mathbf{M}$ of the noisy repeated game is thus calculated as the matrix in Fig. 2. We can see from this transition matrix, although it becomes more complex, it is still a stochastic matrix.

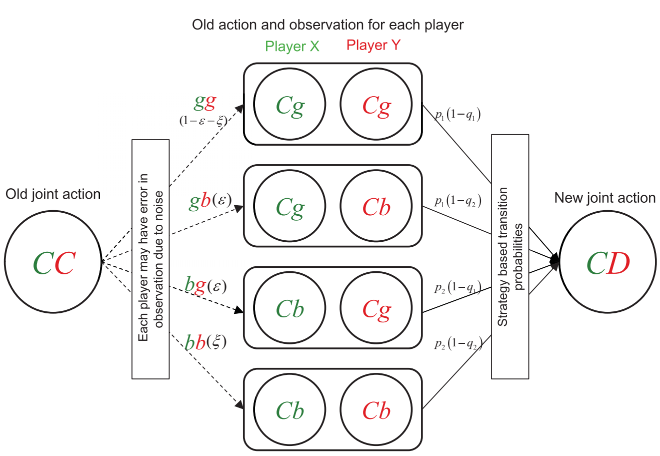

FIG. 1. (Color online) Illustration of the transition from state (joint action) CC to CD. The green color shows the real action and observation of player X while the red color depicts that of player Y. The big nodes denote the action profile, which is the real states of the game. The small nodes denote the combination of one player's action and observation, which are one player's private outcomes.

$$
\begin{array}{c} \\ CC \\ \\ CD \\ \\ DC \\ \\ DD \end{array}
\begin{array}{c}
\begin{array}{cccc} CC & CD & DC & DD \end{array}\\
\begin{bmatrix}
\begin{pmatrix} \tau p_1 q_1 \\ +\varepsilon p_1 q_2 \\ +\varepsilon p_2 q_1 \\ +\xi p_2 q_2 \end{pmatrix} &
\begin{pmatrix} \tau p_1(1-q_1) \\ +\varepsilon p_1(1-q_2) \\ +\varepsilon p_2(1-q_1) \\ +\xi p_2(1-q_2) \end{pmatrix} &
\begin{pmatrix} \tau(1-p_1)q_1 \\ +\varepsilon(1-p_1)q_2 \\ +\varepsilon(1-p_2)q_1 \\ +\xi(1-p_2)q_2 \end{pmatrix} &
\begin{pmatrix} \tau(1-p_1)(1-q_1) \\ +\varepsilon(1-p_1)(1-q_2) \\ +\varepsilon(1-p_2)(1-q_1) \\ +\xi(1-p_2)(1-q_2) \end{pmatrix} \\
\begin{pmatrix} \varepsilon p_1 q_3 \\ +\xi p_1 q_4 \\ +\tau p_2 q_3 \\ +\varepsilon p_2 q_4 \end{pmatrix} &
\begin{pmatrix} \varepsilon p_1(1-q_3) \\ +\xi p_1(1-q_4) \\ +\tau p_2(1-q_3) \\ +\varepsilon p_2(1-q_4) \end{pmatrix} &
\begin{pmatrix} \varepsilon(1-p_1)q_3 \\ +\xi(1-p_1)q_4 \\ +\tau(1-p_2)q_3 \\ +\varepsilon(1-p_2)q_4 \end{pmatrix} &
\begin{pmatrix} \varepsilon(1-p_1)(1-q_3) \\ +\xi(1-p_1)(1-q_4) \\ +\tau(1-p_2)(1-q_3) \\ +\varepsilon(1-p_2)(1-q_4) \end{pmatrix} \\
\begin{pmatrix} \varepsilon p_3 q_1 \\ +\tau p_3 q_2 \\ +\xi p_4 q_1 \\ +\varepsilon p_4 q_2 \end{pmatrix} &
\begin{pmatrix} \varepsilon p_3(1-q_1) \\ +\tau p_3(1-q_2) \\ +\xi p_4(1-q_1) \\ +\varepsilon p_4(1-q_2) \end{pmatrix} &
\begin{pmatrix} \varepsilon(1-p_3)q_1 \\ +\tau(1-p_3)q_2 \\ +\xi(1-p_4)q_1 \\ +\varepsilon(1-p_4)q_2 \end{pmatrix} &
\begin{pmatrix} \varepsilon(1-p_3)(1-q_1) \\ +\tau(1-p_3)(1-q_2) \\ +\xi(1-p_4)(1-q_1) \\ +\varepsilon(1-p_4)(1-q_2) \end{pmatrix} \\
\begin{pmatrix} \xi p_3 q_3 \\ +\varepsilon p_3 q_4 \\ +\varepsilon p_4 q_3 \\ +\tau p_4 q_4 \end{pmatrix} &
\begin{pmatrix} \xi p_3(1-q_3) \\ +\varepsilon p_3(1-q_4) \\ +\varepsilon p_4(1-q_3) \\ +\tau p_4(1-q_4) \end{pmatrix} &
\begin{pmatrix} \xi(1-p_3)q_3 \\ +\varepsilon(1-p_3)q_4 \\ +\varepsilon(1-p_4)q_3 \\ +\tau(1-p_4)q_4 \end{pmatrix} &
\begin{pmatrix} \xi(1-p_3)(1-q_3) \\ +\varepsilon(1-p_3)(1-q_4) \\ +\varepsilon(1-p_4)(1-q_3) \\ +\tau(1-p_4)(1-q_4) \end{pmatrix}
\end{bmatrix}
\end{array}
$$

FIG. 2. Transition matrix of a noisy repeated game.

III. ZD STRATEGIES UNDER NOISE

Let $\mathbf{u}^t$ be the probability distribution over the game's state space $\{CC,CD,DC,DD\}$ at stage t. The probability distributions follow the transition rule such that $\mathbf{u}^{t+1} = \mathbf{u}^t \times \mathbf{M}$. The *stationary distribution* for $\mathbf{M}$ is a vector $\mathbf{v}$ such that $\mathbf{v^T M} = \mathbf{v^T}$. Introducing $\mathbf{M}' = \mathbf{M} - \mathbf{I}$ into the above equation yields $\mathbf{v}^T\mathbf{M}' = \mathbf{0}$. According to Cramer's rule, for any matrix $\mathbf{M}'$ and its adjugate matrix $\mathrm{Adj}(\mathbf{M}')$, the equation $\mathrm{Adj}(\mathbf{M}')\mathbf{M}' = \mathbf{0}$ holds. Therefore from these two equations we know that every row of $\mathrm{Adj}(\mathbf{M}')$ is proportional to the stationary distribution vector $\mathbf{v}$. Changing the last column of $\mathbf{M}'$ into X's stage payoff vector (R_E,S_E,T_E,P_E), we get a new matrix $\tilde{\mathbf{M}}$. Then, using Laplace expansion on the last column of $\tilde{\mathbf{M}}$, we have $\det(\tilde{\mathbf{M}}) = R_E N_1 + S_E N_2 + T_E N_3 + P_E N_4$. The variables N_1, N_2, N_3, and N_4 are just the minors corresponding to R_E, S_E, T_E, and P_E in the last column of $\tilde{\mathbf{M}}$, respectively. The fourth row of $\mathrm{Adj}(\tilde{\mathbf{M}})$ is calculated from the first three columns of $\tilde{\mathbf{M}}$ and is always proportional to $\mathbf{v}$. Therefore X's expected payoff can be calculated by using $\det(\tilde{\mathbf{M}})$. Adding the first column into the second and the third columns gives us a new form of this determinant as in Eq. (2),

$$
\det(\tilde{\mathbf{M}}) = \begin{vmatrix} \cdots & \mu p_1 + \eta p_2 - 1 & \mu q_1 + \eta q_2 - 1 & R_E \\ \cdots & \eta p_1 + \mu p_2 - 1 & \mu q_3 + \eta q_4 & S_E \\ \cdots & \mu p_3 + \eta p_4 & \eta q_1 + \mu q_2 - 1 & T_E \\ \cdots & \eta p_3 + \mu p_4 & \eta q_3 + \mu q_4 & P_E \end{vmatrix}. \quad (2)
$$

In this determinant, $\mu = 1-\varepsilon-\xi$ and $\eta = \varepsilon+\xi$. The first columns is omitted because we only need to analyze the relationship within the last three columns. More importantly, we can see that in this determinant, the second column is solely controlled by X and the third column is solely controlled by Y. We record this new format of determinant as $D(\mathbf{p},\mathbf{q},\mathbf{U_X})$. Then player X's normalized payoff score under stationary state is derived as

$$s_X = \frac{\mathbf{v}\cdot\mathbf{U}_X}{\mathbf{v}\cdot\mathbf{1}} = \frac{D(\mathbf{p},\mathbf{q},\mathbf{U}_X)}{D(\mathbf{p},\mathbf{q},\mathbf{1})}. \tag{3}$$

Similarly, replacing the last column of $\det(\tilde{\mathbf{M}})$ by player Y's stage expected payoff vector, player Y's normalized payoff score is

$$s_Y = \frac{\mathbf{v}\cdot\mathbf{U}_Y}{\mathbf{v}\cdot\mathbf{1}} = \frac{D(\mathbf{p},\mathbf{q},\mathbf{U}_Y)}{D(\mathbf{p},\mathbf{q},\mathbf{1})}. \tag{4}$$

A linear combination of these two scores with coefficients α, β, and γ gives us

$$\alpha s_X + \beta s_Y + \gamma = \frac{D(\mathbf{p},\mathbf{q},\alpha\mathbf{U}_X + \beta\mathbf{U}_Y + \gamma\mathbf{U}_Z)}{D(\mathbf{p},\mathbf{q},\mathbf{1})}. \tag{5}$$

If player X can set his strategy $\mathbf{p}$ delicately and make the second column of this determinant satisfy $\tilde{\mathbf{p}} = \alpha\mathbf{U}_X + \beta\mathbf{U}_Y + \gamma\mathbf{U}_Z$, then the determinant's value $D(\mathbf{p},\mathbf{q},\alpha\mathbf{U}_X + \beta\mathbf{U}_Y + \gamma\mathbf{U}_Z) = 0$, which indicates that X can unilaterally establish a linear relationship between X's and Y's payoff scores, such that $\alpha s_X + \beta s_Y + \gamma = 0$. Such a linear relationship also requires a feasible solution to the following linear equation set:

$$\begin{aligned}
\mu p_1 + \eta p_2 - 1 &= \alpha R_E + \beta R_E + \gamma,\\
\eta p_1 + \mu p_2 - 1 &= \alpha S_E + \beta T_E + \gamma,\\
\mu p_3 + \eta p_4 &= \alpha T_E + \beta S_E + \gamma,\\
\eta p_3 + \mu p_4 &= \alpha P_E + \beta P_E + \gamma.
\end{aligned} \tag{6}$$

If this system of linear equations has feasible solutions, then it will be possible for player X to adjust p_1, p_2, p_3, and p_4 properly to form a linear relationship between his and the opponent's payoffs. Since the above unilateral control strategy is realized by setting a determinant to zero, we call this the *zero-determinant strategy under noise* (NZD strategy for short). Note that when there is no noise (i.e., $\varepsilon = 0$ and $\xi = 0$), the NZD strategy degenerates to the original ZD strategy [2].

IV. PINNING UNDER UNCERTAINTY

One specialization of ZD strategies can unilaterally set the opponent's payoff to a deterministic value [2]. Similar strategies were earlier found by Boerlijst, Nowak, and Sigmund [28]. We call such strategies the pinning strategies. Even in noisy environments, an NZD strategy can establish a pinning property, although the conditions are more strict. If player X chooses proper p_1, p_2, p_3, and p_4, such that $\tilde{\mathbf{p}} = \beta\mathbf{U}_Y + \gamma\mathbf{1}$ (set $\alpha = 0$), then the following linear equation without player X's payoff involved can be formed:

$$\beta s_Y + \gamma = 0. \tag{7}$$

The above $\tilde{\mathbf{p}}$ leads to the following system of linear equations, which depicts the constrains for the pinning strategies under noise:

$$\begin{aligned}
\mu p_1 + \eta p_2 - 1 &= \beta R_E + \gamma,\\
\eta p_1 + \mu p_2 - 1 &= \beta T_E + \gamma,\\
\mu p_3 + \eta p_4 &= \beta S_E + \gamma,\\
\eta p_3 + \mu p_4 &= \beta P_E + \gamma.
\end{aligned} \tag{8}$$

From these four equations we have $\beta = \frac{(\mu-\eta)(p_1-p_2)}{R_E-T_E}$ and $\gamma = p_1 - 1 + \beta\frac{\eta T_E - \mu R_E}{\mu-\eta}$. There are six variables (p_1, p_2, p_3, p_4, β, and γ) in four equations, so we have only two independent free variables. Let p_1 and p_4 be these two variables, and then p_2 and p_3 can be rewritten as

$$\begin{aligned}
p_2 &= \frac{p_1[\mu(T_E - P_E) + \eta(S_E - R_E)] - (1+p_4)(T_E - R_E)}{\mu(R_E - P_E) + \eta(S_E - T_E)},\\
p_2 &= \frac{(1-p_1)(P_E - S_E) + p_4[\mu(R_E - S_E) + \eta(P_E - T_E)]}{\mu(R_E - P_E) + \eta(S_E - T_E)}.
\end{aligned} \tag{9}$$

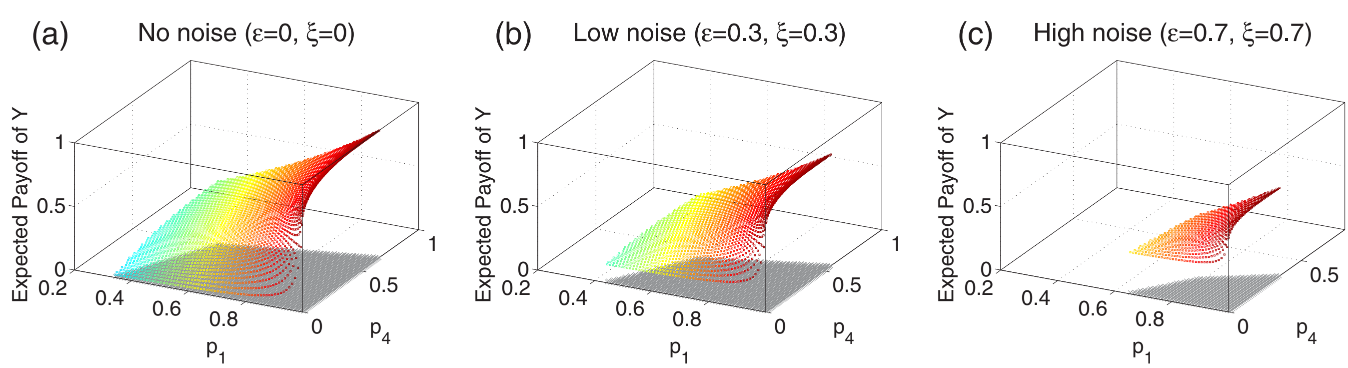

FIG. 3. (Color online) Feasible region of pinning strategies and the corresponding pinned payoffs of player Y under different noises. In each subfigure, the shaded area on the p_1-p_4 plane illustrates the feasible region of pinning strategies. The corresponding pinned payoffs are shown as points on the colored surface. The stage game payoffs are calculated by using $G = 0.5$ and $L = 0.5$, thus realized stage payoffs are $u_i(C,g) = 1, u_i(C,b) = -0.5, u_i(D,g) = 1.5$ and $u_i(D,b) = 0$. The feasible region of pinning strategies as well as the range of pinned payoffs shrink as the noise strength increases. In (a), the game has no noise, thus the expected stage payoffs are $R_E = 1$, $S_E = -0.5$, $T_E = 1.5$, and $P_E = 0$. In (b) the game has low noise and the expected stage payoffs are $R_E = 0.91$, $S_E = -0.41$, $T_E = 1.4$, and $P_E = 0.09$. In (c) there is high noise and the expected stage payoffs are $R_E = 0.79$, $S_E = -0.29$, $T_E = 1.29$, and $P_E = 0.21$.

If and only if the NZD player generates his p_2 and p_3 by following the above formulas, he can pin the expected payoff of the opponent. Representing both β and γ by p_1 and p_4 and substituting them back into Eq. (7), we finally get the opponent's payoff, as

$$s_Y = \frac{(1-p_1)(\mu P_E - \eta S_E) + p_4(\mu R_E - \eta T_E)}{(1-p_1+p_4)(\mu-\eta)}. \quad (10)$$

It is worth noting that, besides the signal distribution, s_Y is only determined by two components in X's strategic vector, which are p_1 and p_4. By inspecting the payoff of Y, we found that in the perfect environment ($\varepsilon = 0$ and $\xi = 0$), player Y's payoff degenerates to $s_Y = \frac{(1-p_1)P_E + p_4 R_E}{(1-p_1)+p_4}$.

From Eqs. (8), the only constrain for the existence of pinning strategies is the probabilistic constrain for p_1, p_2, p_3, and p_4 (i.e., $0 \leqslant p_i \leqslant 1$). We numerically checked the feasible region and the corresponding pinned payoffs of Y, with noise strength ranging from no noise to very strong noise. Since p_2 and p_3 can be represented by p_1 and p_4, we only show the feasible region strategies in the p_1-p_4 plane. As shown in Fig. 3(a), the pinned payoff under the perfect environment arches across whole expected payoff space, ranging from P_E to R_E. However, as the noise is introduced, on the one hand, the feasible region for pinning strategies shrinks, which indicates the noise brings additional constrains for establishing NZD strategies. On the other hand, the range of the pinned payoff also narrows, showing that the NZD player's power of payoff control will be weakened by the noise. In Fig. 3(b), when a weak noise is introduced, the minimum pinned payoff is higher than P_E and the maximum pinned payoff is lower than R_E, and as shown in Fig. 3(c), with the noise strength, the range of the pinned payoff continuously reduces to a very narrow one.

V. EXTORTION UNDER UNCERTAINTY

An NZD strategy in Eqs. (6) can be equivalently rewritten as

$$\tilde{\mathbf{p}} = \varphi[(\mathbf{U}_X - l\mathbf{1}) - \chi(\mathbf{U}_Y - l\mathbf{1})], \quad (11)$$

where φ, χ, and l are free parameters. The only usage of φ is to ensure the probabilities to locate in [0,1]. It is worth noting that if $l \leqslant P_E$, then the probability constraints cannot be satisfied and NZD strategies do not exist. Thus we only need to investigate different cases when $l \geqslant P_E$. In the case (i) $\chi \to \infty$, $\mathbf{p}$ is a pinning strategy. In the case (ii) $\chi > 1$ and $l \geqslant P_E$, player X can ensure that, when player Y tries to increase his payoff, he will increase X's even more, and X's increase of payoff exceeds that of Y by a fixed percentage χ. In addition, Y can only maximize his payoff by fully cooperating ($\mathbf{q} = \mathbf{1}$). Therefore, if player X chooses a $\mathbf{p}$ with $\chi > 1$, then X can always extort Y since Y's effort will benefit X more than himself. In the case (iii) $\chi > 1$ and $l = P_E$, player X not only ensures his payoff increment is χ-fold of Y's but also guarantees that his absolute payoff is higher than Y's and, consequently, dominates in the game.

We distinguish the second and the third cases and call the former the *contingent extortion* strategy and the latter the *dominant extortion* strategy. A dominant extortion is a secure strategy, in the sense that the NZD player using such a strategy not only grabs the achievement of the opponent but also always outperforms the opponent. On the contrary, using a contingent extortion, although the NZD player can guarantee a higher increment of payoff, he cannot secure that he can always outperform the opponent, which means the NZD players' outcome of the game is contingent. The word "contingent" captures such uncertainty that the NZD player may or may not dominate in payoff, even though he can always obtain a higher payoff increment than the opponent. It is worth noting that any dominant extortion is a special and most stringent case of contingent extortions, and any contingent extortion is an instantiation of NZD strategies. Essentially, the uncertainty of extortion is quantitatively affected by the parameter l, which can be seen as the *baseline* of extortion. In noisy games, the contingency or uncertainty is caused by the perception or observation errors, thus, in such games, l can be defined as a function of noise structure.

Although the dominant extortion strategies are found widely existing in games without noise [2], we prove that in noisy repeated games, the dominant extortion strategies do not exist. To enforce a dominant extortion strategy, according to Eq. (5), the following equation set is required to be satisfied when $l = P_E$:

$$\begin{aligned} \mu p_1 + \eta p_2 - 1 &= \varphi[(R_E - l) - \chi(R_E - l)],\\ \eta p_1 + \eta p_2 - 1 &= \varphi[(S_E - l) - \chi(T_E - l)],\\ \mu p_3 + \eta p_4 &= \varphi[(T_E - l) - \chi(S_E - l)],\\ \eta p_3 + \mu p_4 &= \varphi[(P_E - l) - \chi(P_E - l)]. \end{aligned} \quad (12)$$

However, when $l = P_E$, the third and the fourth equations cannot be satisfied simultaneously. Intuitively, the lack of the dominant extortion strategy in noisy repeated games is because the errors introduce stochasticity and uncertainty into the payoffs and, consequently, have a negative impact on the accuracy of player X's payoff-based strategy setting. Therefore, the NZD player faces a fundamental trade-off between the payoff control ability and the payoff dominance. Such a trade-off is similar to the relationship between risk dominance and payoff dominance, which has been discussed in pioneering works by Harsanyi and Selten [29]. Thus, in a noisy environment, to regain the payoff control ability, the extortioner needs to relax the extortion baseline from P_E to $P_E + \Delta$, which, on the contrary, increases the risk for him to lose in payoff. We represent the contingent extortion strategy as a (χ,Δ)-extortion strategy, where χ defines the extortion rate while $\Delta = l - P_E$ defines the distance between the weak and dominant extortion strategies that can be considered as the *generosity* [8]. When Δ is small, it is still very likely (though not necessarily so) for player X to always get higher payoffs than player Y; however, it will be difficult for him to establish an extortion on player Y's payoff. A larger Δ indicates that player X offers more opportunity for the opponent to win in payoff but correspondingly obtains higher possibility for himself to control the opponent's payoff. Therefore, in order to realize a payoff control while reducing the risk of losing, it is of great importance for the NZD player to design his strategy with a proper extortion ratio χ and a sufficiently small distance Δ.

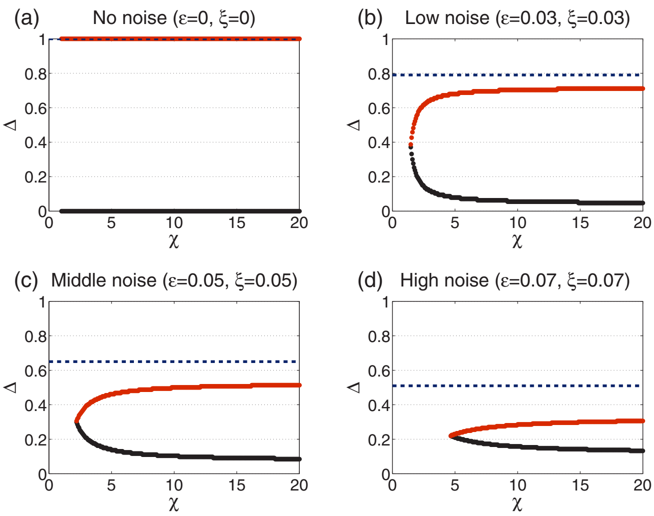

FIG. 4. (Color online) Feasible regions for contingent extortion strategies under different noise strengths. The black curves and red curves depict the lower bounds and upper bounds of Δ versus χ, respectively. The blue dashed lines show the values of $R_E - P_E$. In all subfigures, player X's realized payoff is set as $u_X(C,g) = 1$, $u_X(C,b) = -0.5$, $u_X(D,g) = 2$, and $u_X(D,b) = 0$. The expected stage payoffs R_E, S_E, T_E, and P_E are calculated by Eq. (1). For the noise-free case (a), the lower bound of Δ is always 0 and the upper bound of Δ is always 1, indicating that NZD strategies always exist for any χ. In the low-noise case (b), the contingent extortion strategies with small χ ($\chi < 1.78$) do not exist, and the feasible range of Δ becomes larger as the increase of χ after it exceeds 1.78. The lower bound approaches a value greater than 0 while the upper bound approaches a value smaller than 0.79. In (c) there is a medium amount of noise, and in (d) there is much more noise. Comparing (a), (b), (c), and (d), it is found that the feasible region of contingent extortion strategies dramatically shrinks with the increase of noise strength.

According to the analysis above, to get a contingent extortion strategy under noise, the following vector equation is required:

$$\tilde{\mathbf{p}} = \varphi\{[\mathbf{U}_X - (P_E + \Delta)\mathbf{1}] - \chi[\mathbf{U}_Y - (P_E + \Delta)\mathbf{1}]\}, \quad (13)$$

which can be expanded to:

$$\begin{aligned} p_1 &= 1 - \varphi\frac{1}{\tau - r}[F_1 - \chi F_2] + \varphi(\chi - 1)\Delta, \\ p_2 &= 1 + \varphi\frac{1}{\tau - r}[J_1 - \chi J_2] + \varphi(\chi - 1)\Delta, \\ p_3 &= \varphi\frac{(\tau + \varepsilon)}{\tau - r}[T_E - P_E - \chi(S_E - P_E)] + \varphi(\chi - 1)\Delta, \\ p_4 &= -\varphi\frac{(\varepsilon + r)}{\tau - r}[T_E - P_E - \chi(S_E - P_E)] + \varphi(\chi - 1)\Delta, \end{aligned} \quad (14)$$

where $F_1 = \mu R_E - \eta S_E - (\mu - \eta)P_E$, $F_2 = \mu R_E - \eta T_E - (\mu - \eta)P_E$, $J_1 = \eta R_E - \mu S_E + (\mu - \eta)P_E$, and $J_2 = \eta R_E - \mu T_E + (\mu - \eta)P_E$. As shown in Fig. 4, we numerically checked the feasible region of contingent extortion strategies by exploring the whole space of Δ versus a different extortion ratio χ. One can see that the distance Δ has both a lower bound and an upper bound, with the former positively correlated with the noise strength and the latter negatively correlated with the noise strength. Combining these two effects, the feasible range of Δ shrinks while the noise becomes stronger. In addition, increasing the lower bound suggests that the NZD player should relax its extortion baseline l and move it farther from P_E as the noises strength increases.

When player X adopts a contingent extortion strategy, the payoffs of players X and Y follow the following linear relationship:

$$s_X - (P_E + \Delta) = \chi[s_Y - (P_E + \Delta)]. \quad (15)$$

Since in the prisoner's dilemma $T_E > R_E > P_E > S_E$, X's payoffs when Y chooses action C (T_E or R_E) are always larger than his payoffs when Y chooses action D (P_E or S_E). The same result holds when player Y mixes his action. Thus whatever strategy X takes, its expected payoff s_X will be maximized when Y fully cooperates ($\mathbf{q} = \mathbf{1}$). When X takes contingent extortion strategy, since s_X and s_Y follow a linear relationship, s_Y will also be maximized when s_X reaches its maximum. Therefore, both s_X and s_Y are maximized when Y fully cooperates. Substituting $q_1 = q_2 = q_3 = q_4 = 1$ into $\det(\tilde{\mathbf{M}})$, the determinant becomes

$$\det(\mathbf{p},\mathbf{1},\mathbf{U}_X) = \begin{vmatrix} 1 - \mu p_1 + \eta p_2 & 0 & 0 & R_E \\ \eta p_1 + \mu p_2 & -1 & 1 & S_E \\ \mu p_3 + \eta p_4 & 0 & 0 & T_E \\ \eta p_3 + \mu p_4 & 0 & 1 & P_E \end{vmatrix}. \quad (16)$$

Making Laplace expansion on the fourth column, we have the normalized payoff for player X as

$$s_X = \frac{\det(\mathbf{p},\mathbf{1},\mathbf{U}_X)}{\det(\mathbf{p},\mathbf{1},\mathbf{1})},$$

which finally leads to

$$\begin{aligned} s_X = {} & \frac{1}{C}\chi[R_E(T_E - S_E) - P_E(T_E - R_E)] \\ & - \frac{1}{C}(\chi - 1)(T_E - R_E)\Delta + \frac{1}{C}P_E(T_E - R_E) \end{aligned} \quad (17)$$

and

$$\begin{aligned} s_Y = {} & \frac{1}{C}\chi P(R_E - S_E) + \frac{1}{C}(\chi - 1)(R_E - S_E)\Delta \\ & + \frac{1}{C}[S_E(P_E - R_E) + P_E(T_E - R_E)], \end{aligned} \quad (18)$$

where $C = (T_E - R_E) + \chi(R_E - S_E) > 0$. For instance, if $(R_E,S_E,T_E,P_E) = (3,0,5,1)$, we have

$$s_X = \frac{2 + 13\chi - 2(\chi - 1)\Delta}{2 + 3\chi}, \quad (19)$$

and, accordingly, the payoff for player Y is

$$s_Y = \frac{12 + 3\chi + 3(\chi - 1)\Delta}{2 + 3\chi}. \quad (20)$$

In a word, on the one hand, the extortion strategies are still feasible in a noisy environment, which indicates it is still possible for the NZD player to ensure that when the opponent tries to improve his payoff, he will improve the NZD player's even more. And the opponent will maximize his own payoff by fully cooperating, where the NZD player's payoff is also maximized. Thus the NZD player can still enforce a contingent extortion on his opponent. However, on the other hand, the uncertainty in the noisy environment has abated the power of extortion, in the sense that the extortioner cannot guarantee his payoff always to be higher than the opponent's and the dominant extortion strategies do not exist. The baseline for contingent extortion strategies should have a distance to P_E, and the lower bound of the distance has a positive correlation with noise strength. Under the same extortion ratio χ, the payoffs for the extortioner and for the opponent under different noise strengths varies. In Eq. (17) we can see s_X may decline as noise strength increases. On the contrary, in Eq. (18), s_Y may increase as noise strength increases. Therefore, under a certain noise strength (which results in a reasonably large distance), it is possible for s_Y to outperform s_X. These indicate that, in noisy environments, when an NZD player wishes to extort the opponent and control the payoffs, there is an increased risk for him to loss in payoff, especially when the noise is strong. Therefore, in a realistic uncertain world, extorting others has the potential to cause damage to oneself.

VI. CONCLUSION AND DISCUSSION

The concept of the ZD strategy has become a promising framework to explore long-run relationships. However, outside of the laboratory, the existence of noise in the environment elevates the complexity of games and payoff-oriented ZD strategy selection in such games deserves more concrete analysis. We established the generalized form of the ZD strategy for noisy games and named it the NZD strategy. We identify three specifications of NZD strategies, namely the pinning strategies, dominant extortion, and contingent extortion. We also studied the conditions, feasible regions, and corresponding payoffs for these strategies. It is found that NZD strategies have high robustness against noise and widely exist in noisy games with reasonable noise strength, although the noise has a negative impact on the existence and performance of NZD strategies. The noise will expose the NZD player to uncertainty and risk; however, it is still possible for him to set the opponent's payoff to a fixed value or to extort the opponent.

The implementation of the NZD strategies relies on the existence of the unique stationary distribution. However, not only the existence of noise but also some special strategies may result in bad circumstances such that the regularity of the Markov matrix cannot be satisfied or the Markov process may not converge to a unique stationary distribution. Thus it is essential to analyze the convergency of the Markov process of the game. This is not only important to ZD or NZD strategies but also a key problem for other topics in repeated games. When multiple stationary distribution exists, the Markov process may have multiple converging states, which belong to different communicating classes. In this case, the expected payoff of each player is strongly affected by the initial state of the game. We conjecture that, in a game with multiple stationary distribution, a generalized NZD strategy whose expected payoff is engaged with initial distribution may still exist. Moreover, the speed for the Markov process to converge is a key factor for the NZD player. The second-largest eigenvalue of a Markov transition matrix is a convenient factor to determine which strategy of the NZD player may lead the game to converge faster. Although the converging speed is not unilaterally determined by the NZD player, he can at least secure himself with a maximized lower boundary of the converging speed.

Another condition for extortion under uncertainty is that the NZD player should know the error distribution. This is a common assumption in the literature about noisy repeated games. The modeling in this paper is also based on such an assumption. In particular, when both players know the error distribution, the long-run relationship and the decision making involve complicated reinforcement learning and statistical inference. In this case, the historical information of the game can be utilized for a player to determine what he is going to do to improve his or long-term payoffs, and this can become increasingly more complex as time goes by [19]. On the one hand, as long as the NZD player is aware of the error distribution, he can efficiently launch the NZD strategies and enforce a linear relationship between the two players' payoffs, regardless of the opponent's strategy. However, the NZD player can make a further step and learn the pattern of the opponent's behavior by utilizing the known error distribution. We have proved that dominant extortion does not exist in noisy games and that the NZD player may have a small risk of losing even he can control the payoffs of the opponent. Learning the pattern of the opponent's strategy will help the NZD player to intelligently alter his strategy under different circumstances and avoid such a risk. On the other hand, even if the opponent knows the error distribution, as long as he does not learn the NZD strategy, it is very likely that he will be distracted by the NZD player. However, if his learning is sufficient and he is capable of identifying the NZD player's strategy, he may take advantage of the risk of the NZD player and accordingly design his strategy, resulting in a winning situation under uncertainty.

Furthermore, the original ZD strategies do not necessarily promote cooperation, since the Markov process does not surely converge to a joint state CC. When the repeated game is played in an imperfect environment, this becomes even more severe. The *generous strategies* [8] not only guarantee a linear relationship between two players' payoffs but also ensure that the mutual cooperation payoff is the maximum payoff to both the ZD player and the opponent. Generosity comes at a cost, but it finally encourages everybody to cooperate. Although the generous strategies are proved to be very robust in the perfect environment, whether it exists and how it performs in the noisy environment still need investigation. In particular, how can we provide a strategy that makes the game always converge to the mutual cooperation state, even when the noise has disturbances during mutual cooperation? Actually, this topic is strongly related to the equilibrium analysis in repeated games with private monitoring, which is the one of the most well-known long-standing open problems in game theory research [19]. The framework of NZD strategies may potentially provides us with another possible direction to tackle this issue.

ACKNOWLEDGMENT

This work was partially supported by the National Natural Science Foundation of China under Grants No. 61473060 and No. 11222543.

[1] J. Mailath and L. Samuelson, *Repeated Games and Reputation* (Oxford University Press, Oxford, 2006).
[2] W. H. Press and F. J. Dyson, Proc. Acad. Natl. Sci. USA **109**, 10409 (2012).
[3] A. J. Stewart and J. B. Plotkin, Proc. Acad. Natl. Sci. USA **109**, 10134 (2012).
[4] B. Hayes, Am. Sci. **100**, 422 (2013).
[5] D. Hao, Z. Rong, and T. Zhou, Chin. Phys. B **23**, 078905 (2014).
[6] L. Roemheld, arXiv:1308.2576 (2013).
[7] E. Akin, arXiv:1211.0969 (2012).
[8] A. J. Stewart and J. B. Plotkin, Proc. Acad. Natl. Sci. USA **110**, 15348 (2013).
[9] C. Hilbe, M. A. Nowak, and A. Traulsen, PLoS ONE **8**, e77886 (2013).
[10] J. Chen and A. Zinger, J. Theor. Biol. **357**, 46 (2014).
[11] L. Pan, D. Hao, Z. Rong, and T. Zhou, arXiv:1402.3542 (2014).
[12] C. Hilbe, B. Wu, A. Traulsen, and M. A. Nowak, Proc. Acad. Natl. Sci. USA **111**, 16425 (2014).
[13] C. Hilbe, T. Rohl, and M. Milinski, Nat. Commun. **5**, 3976 (2014).
[14] C. Adami and A. Hintze, Nat. Commun. **4**, 2193 (2013).
[15] C. Hilbe, M. A. Nowak, and K. Sigmund, Proc. Acad. Natl. Sci. USA **110**, 6913 (2013).
[16] A. Szolnoki and M. Perc, Phys. Rev. E **89**, 022804 (2014).
[17] Z. X. Wu and Z. Rong, Phys. Rev. E **90**, 062102 (2014).
[18] A. Szolnoki and M. Perc, Sci. Rep. **4**, 5496 (2014).
[19] M. Kandori, J. Econ. Theor. **102**, 1 (2002).
[20] M. A. Nowak, K. Sigmund, and E. El-Sedy, J. Math. Biol. **33**, 703 (1995).
[21] D. Fudenberg, G. R. David, and D. Anna, Am. Econ. Rev. **102**, 720 (2012).
[22] D. Fudenberg, D. Levine and E. Maskin, J. Econometric Soc. **62**, 997 (1994).
[23] T. Sekiguchi, J. Econ. Theor. **76**, 345 (1997).
[24] M. Barlo, C. Guilherme, and S. Hamid, J. Econ. Theor. **144**, 312 (2009).
[25] G. J. Mailath and M. Stephen, J. Econ. Theor. **102**, 189 (2002).
[26] G. J. Mailath and O. Wojciech, Gam. Econ. Behav. **71**, 174 (2011).
[27] B. B. Murdock, J. Exp. Psy. **64**, 482 (1962).
[28] M. C. Boerlijst, M. A. Nowak, and K. Sigmund, Am. Math. Mon. **104**, 303 (1997).
[29] J. C. Harsanyi and R. Selten, *A General Theory of Equilibrium Selection in Games* (MIT Press, Cambridge, MA, 1988).