TOPICAL REVIEW — Statistical physics and complex systems

Zero-determinant strategy: An underway revolution in game theory*

Hao Dong(郝 东), Rong Zhi-Hai(荣智海), and Zhou Tao(周 涛)†

Evolutionary Game Institute, Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 611731, *China*

(Received 3 March 2014; revised manuscript received 20 May 2014; published online 13 June 2014)

Repeated games describe situations where players interact with each other in a dynamic pattern and make decisions according to outcomes of previous stage games. Very recently, Press and Dyson have revealed a new class of zero-determinant (ZD) strategies for the repeated games, which can enforce a fixed linear relationship between expected payoffs of two players, indicating that a smart player can control her unwitting co-player's payoff in a unilateral way [*Proc. Acad. Natl. Sci. USA* **109**, 10409 (2012)]. The theory of ZD strategies provides a novel viewpoint to depict interactions among players, and fundamentally changes the research paradigm of game theory. In this brief survey, we first introduce the mathematical framework of ZD strategies, and review the properties and constrains of two specifications of ZD strategies, called pinning strategies and extortion strategies. Then we review some representative research progresses, including robustness analysis, cooperative ZD strategy analysis, and evolutionary stability analysis. Finally, we discuss some significant extensions to ZD strategies, including the multi-player ZD strategies, and ZD strategies under noise. Challenges in related research fields are also listed.

Keywords: zero-determinant strategy, prisoner's dilemma, evolution of cooperation, repeated game

PACS: 89.75.Fb, 02.50.Le, 89.65.–s **DOI:** 10.1088/1674-1056/23/7/078905

1. Introduction

Cooperative and altruistic behaviors seem at odds in the Darwinian principle, which is based on the concept of competition. However, this concept does not explain the enormous amount of cooperations that are visible in nature.[1–6] Scientists have used game theory to study the cooperation and competition in both human society and evolutionary biology. Cooperative dilemmas are situations in which collective interests are at odds with private interests. In a one-shot interaction between players, each player always chooses selfish strategy ignoring the opponent's choice, the cooperative dilemma thus takes place and makes the game get ugly. The collapse of cooperation can be regarded as a simple model of "the tragedy of the commons".[7] To avoid the tragedy, one valid approach is to build up a long-run relationship, in other words, to make the game repeat. The repeated game[8] captures the idea that a player needs to take into account the impacts of her actions on the future feedbacks of other players, which has strong relevance with reputation and social norm.[9] Among all the repeated games, the iterated prisoner's dilemma (IPD)[6] has a long history as a model for the study of evolution of cooperation.

Very recently, a class of zero-determinant (ZD) strategies, discovered by Press and Dyson,[10] have attracted considerable attention. By using ZD strategies in a repeated game, a player can unilaterally pin her opponent's expected payoff to her desirable value, or extort her opponent by enforcing a linear relationship between her and the opponent's payoffs. Thus, ZD strategies provide the player with a strong unilateral control in games. In the framework of ZD strategy, there are dozens of ingenious ideas and untraditional approaches for analyzing not only prisoner's dilemma but also bi-matrix games, which dramatically expand our understanding of the stochastic process, the mutual benefit, the cooperation incentive, and even the optimal control in the repeated games.

The impact of ZD strategies is multifold. For game theory itself, the discovery and formal definition of unilateral controlling strategy have already built a milestone.[11,12] Following Press and Dyson's work, future research on repeated games will surely be framed in terms of ZD strategies. The related topics cover the cooperation enforcement,[13] the equilibrium analysis[13] and the robustness analysis based on the framework of ZD strategies.[14] On the other hand, there have already sprung up dozens of works studying the impact of ZD strategies in the evolutionary games, such as analyzing their evolutionary stability and simulating their performance under different environments.[15–18] Furthermore, after the discovery of ZD strategies, several research articles worked on the taxonomy of different strategies. They have clearly revealed the underlying relationships between the pinning strategies, the ex-

*Project supported by the National Natural Science Foundation of China (Grant Nos. 61004098 and 11222543), the Program for New Century Excellent Talents in Universities of China (Grant No. NCET-11-0070), the Special Project of Youth Science and Technology Innovation Research Team of Sichuan Province, China (Grant No. 2013TD0006), and the Research Foundation of UESTC and Scholars Program of Hong Kong (Grant No. G-YZ4D).

†Corresponding author. E-mail: zhutou@ustc.edu

© 2014 Chinese Physical Society and IOP Publishing Ltd

tortion strategies and the classic strategies such as tit-for-tat (TFT), win-stay, lose shift (WSLS) and so on.[18,19]

2. Starting from two-player prisoner's dilemma

To study the interactive decision-making problems, von Neumann and Morgenstern[20] invented the mathematical framework of games, which provide a powerful tool to understand the emergence and maintenance of cooperation. Nash[21] introduced the celebrated solution concept "Nash equilibrium", which discloses a state in games such that no player can increase her own payoff by altering strategy unilaterally. In order to characterize the interactions among individuals, many game models are proposed, among which the prisoner's dilemma (PD) is most well known. In the PD game, there are two players X and Y, each of which may either cooperate (C) or defect (D). If both of them cooperate, each earns a reward payoff R. If a cooperator meets a defector, the defector gains a large temptation payoff T, while the cooperator gets a small sucker payoff S. If both of them defect, each gains a mutual punishment payoff P. The PD games have two constrains: (i) $T > R > P > S$, which guarantees that mutual defection is one of the Nash equilibriums; (ii) $2R > T + S$ ensures that there is some incentive for cooperation. In such a game, the selfish individuals fall into a dilemma that both of them receive a lower payoff as defectors, although they can obtain a higher payoff through cooperating with each other.

If the PD game is played only once, the only strong Nash equilibrium is mutual defection and it is difficult to maintain cooperation. However, the situation may be different if the players interact repeatedly. Axelrod and Hamilton[1] investigated different strategies in a repeated PD game with a payoff vector $(T,R,P,S) = (5,3,1,0)$. Through two rounds of computer tournaments, they found that the winning strategy is always the simplest, named "tit-for-tat" (TFT). A TFT player firstly cooperates with her opponent, then imitates the opponent's previous action. The TFT strategy has three good features. First of all, it is a nice strategy, which means it never defects first. Secondly, it can punish a defector. Thirdly, it forgives an opponent who ever defected but cooperated in the recent round. However, one major shortcoming of TFT is it cannot correct any mistakes in the noisy environment. Nowak and Sigmund proposed two strategies, the "generous tit-for-tat" (GTFT)[22] and the "win-stay, lost-shift" (WSLS).[23] The GTFT player not only cooperates with a cooperator, but also cooperates with a defector with a certain probability. The WSLS player repeats the previous action if the resulting payoff has met her aspiration level and changes to another action otherwise. The strategies TFT, GTFT and WSLS are all memory-one strategies, which means the player's choice in the current round is only conditioned on the actions of players in the last round.

In the real world, players can only make decisions according to limited information. Smith[24] proposed the evolutionary game theory, in which it is not necessary to assume the rationality of players, instead, players can update their strategies through learning the successful behaviors from others. Smith introduced the evolutionarily stable strategy (ESS), which is a strategy that if it is adopted by a population in a given environment, it cannot be invaded by any other strategies that are initially rare. ESS is a stable Nash equilibrium if the size of population approaches infinite. Nowak *et al.*[25] proposed the evolutionarily stable strategy for finite population. They found that the size of population plays a crucial role. Within a finite population, natural selection favors cooperative strategies instead of defective strategies.

3. Zero-determinant strategies in iterated prisoner's dilemma

3.1. Framework

Press and Dyson[10] proved that, if a fixed game is played repeatedly, the shortest memory player sets the rule of the game, which means the long-memory strategies have no advantages over the memory-one strategies. In a repeated game, the outcome of each stage game can be represented as $xy \in \{CC,CD,DC,DD\}$, where C and D denote cooperation and defection, respectively. X's mixed strategy is a vector $\mathbf{p} = (p_1,p_2,p_3,p_4)$, showing the probability to cooperate corresponding to the above outcomes. Analogously, Y's strategy is $\mathbf{q} = (q_1,q_2,q_3,q_4)$, a probability vector conditioned on the outcomes $yx \in \{CC,CD,DC,DD\}$. For example, the strategy TFT is denoted as a vector $(1,0,1,0)$.

A repeated game is usually characterized by a Markov process, in which the states are the outcomes in stage games. The state transition of the Markov process is fully determined by the two players' probabilistic strategies $\mathbf{p}$ and $\mathbf{q}$, and the transition matrix $\mathbf{M}(\mathbf{p},\mathbf{q})$ is defined as

$$\begin{array}{c} \\ cc \\ cd \\ dc \\ dd \end{array} \begin{array}{c} \begin{array}{cccc} cc & cd & dc & dd \end{array} \\ \begin{bmatrix} p_1q_1 & p_1(1-q_1) & (1-p_1)q_1 & (1-p_1)(1-q_1) \\ p_2q_3 & p_2(1-q_3) & (1-p_2)q_3 & (1-p_2)(1-q_3) \\ p_3q_2 & p_3(1-q_2) & (1-p_3)q_2 & (1-p_3)(1-q_2) \\ p_4q_4 & p_4(1-q_4) & (1-p_4)q_4 & (1-p_4)(1-q_4) \end{bmatrix} \end{array}. \quad (1)$$

Denote $\mathbf{v}$ as the stationary probability vector of $\mathbf{M}$, such that $\mathbf{v}\cdot\mathbf{M} = \mathbf{v}$, then the inner product $\mathbf{v}\cdot\mathbf{S}_X$ yields player X's expected payoff in the stationary state, where $\mathbf{S}_X = (R,S,T,P)$ is player X's payoff vector. Press and Dyson[10] proved that $\mathbf{v}\cdot\mathbf{S}_X$ equals the determinant of a matrix which is obtained via replacing the last column of $\mathbf{M}-\mathbf{I}$ by $\mathbf{S}_X$. Denote this determinant as $\mathscr{D}(\mathbf{p},\mathbf{q},\mathbf{S}_X)$, then player X's normalized expected payoff is $E_X = \mathscr{D}(\mathbf{p},\mathbf{q},\mathbf{S}_X)/\mathscr{D}(\mathbf{p},\mathbf{q},\mathbf{1})$, where $\mathbf{1}$ is an all-ones vector introduced for normalization. In the determi-

nant, the second column is solely determined by the strategy of player X and the third column is solely determined by the strategy of player Y. We record the second column as $\tilde{\mathbf{p}} = (-1+p_1, -1+p_2, p_3, p_4)$ and the third column as $\tilde{\mathbf{q}} = (-1+q_1, q_3, -1+q_2, q_4)$. Note that the same result also holds for $\mathbf{v}\cdot\mathbf{S}_Y$, where $\mathbf{S}_Y = (R,T,S,P)$ is the payoff vector for player Y. Thus, player Y's normalized expected payoff is $E_Y = \mathscr{D}(\mathbf{p},\mathbf{q},\mathbf{S}_Y)/\mathscr{D}(\mathbf{p},\mathbf{q},\mathbf{1})$. Performing linear combination of E_X and E_Y yields

$$\alpha E_X + \beta E_Y + \gamma = \frac{\mathscr{D}(\mathbf{p},\mathbf{q},\alpha\mathbf{S}_X+\beta\mathbf{S}_Y+\gamma\mathbf{1})}{\mathscr{D}(\mathbf{p},\mathbf{q},\mathbf{1})}. \tag{2}$$

In Eq. (2), if player X sets $\mathbf{p}$ properly and makes the second column $\tilde{\mathbf{p}}$ equal $\alpha\mathbf{S}_X+\beta\mathbf{S}_Y+\gamma\mathbf{1}$, then she can unilaterally make the second column and the last column in the matrix identical, and consequently, make the determinant vanish. Then a linear relationship between two players' expected payoffs will be established. Such linear relationships potentially exist not only in PD, but in every two-player games. Since the above strategies are derived from setting the determinant of a matrix to be zero, this kind of strategies are named the zero-determinant (ZD) strategies.

The discovery of ZD strategies changes our understanding of game theory. Observing the determinant-based payoffs, we may find they are very concise and operable. Based on the above framework, in the following subsections, we will review two kinds of important specializations of ZD strategies: (1) strategies that fix the opponent's expected payoff at a certain value, which are called pinning strategies; (2) strategies that guarantee a player with a larger surplus than her opponent, which are called extortion strategies.

3.2. Pinning strategies

It is found by surprise that one player can choose a subclass of ZD strategies that can determine her opponent's expected payoffs, regardless of the opponent's choice. We call such strategies the pinning strategies. A pinning strategy can be realized by choosing $\mathbf{p}$ that satisfies $\tilde{\mathbf{p}} = \beta\mathbf{S}_Y + \gamma\mathbf{1}$ (i.e., $\alpha = 0$). Following this strategy, equation (2) is reduced to

$$\beta E_Y + \gamma = \frac{\mathscr{D}(\beta\mathbf{S}_Y+\gamma\mathbf{1},\ \mathbf{q},\ \beta\mathbf{S}_Y+\gamma\mathbf{1})}{\mathscr{D}(\beta\mathbf{S}_Y+\gamma\mathbf{1},\mathbf{q},\mathbf{1})} = 0,$$

and player Y's payoff is fixed to value $-\gamma/\beta$. Since both β and γ can be represented by the components in $\mathbf{p}$, the expected payoff of player Y is unilaterally determined by player X' strategy, but does not depend on its own strategy anymore.

After combination and reduction, Y's expected payoff can be represented by only p_1 and p_4, which are the probabilities for cooperation when the previous outcomes are mutual cooperation (CC) and mutual defection (DD), respectively. Thus, player Y's expected payoff E_Y is a weighted average of P and R with weights $1-p_1$ and p_4, such that

$$E_Y = \frac{(1-p_1)P + p_4 R}{(1-p_1)+p_4}. \tag{3}$$

By inspecting the boundary value of E_Y, we see that player X can pin the opponent's payoff to any value in the range from the mutual defection payoff P to the mutual cooperation payoff R. Figure 1(a) illustrates the payoff pairs with player X adopting a pinning strategy while player Y choosing randomized strategies. The sample points fall into a straight line with slope zero, which we called the pinning line. The vertical position of this pinning line is totally determined by player X's strategic probabilities p_1 and p_4, which reflects player X's toughness.

For a given PD payoff vector (R, S, T, P), pinning strategies only exist when p_1 is sufficiently large and p_4 is sufficiently small. A smaller p_1 or a smaller p_4 indicates that player X is tougher and inclined to give a punishment. Under this situation, the vertical position of the pinning line is approaching the mutual punishment value P. Correspondingly, larger p_1 or p_4 makes the payoff pairs closer to the mutual

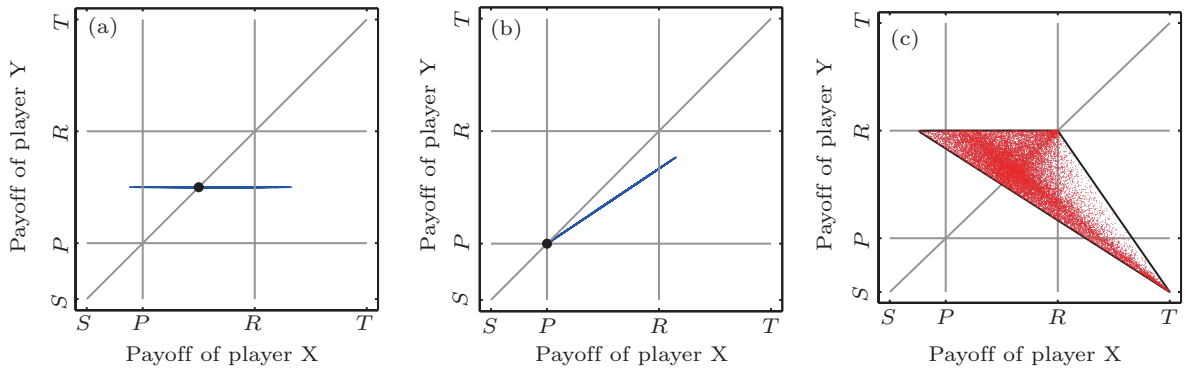


Fig. 1. Payoff pairs when player X uses different strategies against a random strategy player Y. (a) Player X uses a pinning strategy with $\beta = -0.2, \gamma = 0.1$, and $\mathbf{p} = (3/7, 0, 5/7, 2/7)$. One can observe that player Y's payoff is pinned to a line with slope zero. However, the payoff of X cannot essentially be fixed to a certain value. (b) An extortion strategy can be formed, if player X sets $\chi = 1.5$ and $\varphi = 0.01$. Under this strategy, since the line of payoff pairs is under the diagonal line, X's payoff surplus over mutual punishment P is always larger than Y. Panels (a) and (b) are in a stark contrast with panel (c), where player X adopts WSLS and the payoff pairs cover a 2-dimensional area.

cooperation value R. Using such pinning strategies, player X achieves a strong control on Y's expected payoff. She has no more needs to react to Y. Moreover, she can even experiment an evolutionary path for Y.

3.3. ZD player cannot fix her own payoff

Press and Dyson[10] proved that, although a pinning strategy player can control her opponent's payoff, she cannot guarantee herself a high payoff. This is because that, if X wishes to set her own payoff, she needs to choose a strategy that satisfies $\tilde{\mathbf{p}} = \alpha \mathbf{S}_X + \gamma \mathbf{1}$. However, the only solution for this equation is the singular strategy $\mathbf{p} = (1,1,0,0)$. Then the state transition matrix $\mathbf{M}$ becomes a block diagonal matrix with two closed communicating classes, corresponding to a Markov process without a unique stationary distribution. That is to say, the stationary state of the game depends on the initial state distribution, thus one player cannot pin the expected payoff of her own to a fixed value.

Although player X cannot fix her payoff, it is encouraged to investigate whether it is still possible for her to establish a region of desirable payoffs. In Fig. 1(a), the straight line of the payoff pairs intersects the diagonal line at one point, where the two players' payoffs are equal. On the left side of this intersection is the region where player X's payoff is smaller than Y's, while on the right side, player X's payoff is larger than Y's. A pinning strategy player can enforce most of the payoff pairs locating on the right side of the diagonal line, where her own payoff is always greater than her opponent's. She can do this by decreasing the probabilities of cooperation, and consequently, push the pinning line downwards. In this case, although player X is tough and successfully suppresses her opponent's payoff to a very limited value, she may also suffer loss since the opponent may have no incentive to cooperate anymore. If player Y is not simply a random strategy player, such a tough pinning strategy may lead to a lose–lose situation and end the game with mutual punishment.

3.4. Extortion strategies

By using another subclass of ZD strategies, one player can guarantee that her own surplus over the mutual punishment payoff P is χ-fold of the opponent's. These are called extortion strategies. An extortion strategy can be realized by setting $\tilde{\mathbf{p}} = \varphi[(\mathbf{S}_X - P\mathbf{1}) - \chi(\mathbf{S}_Y - P\mathbf{1})]$, where $\chi \geq 1$ represents the extortion ratio and φ ensures that the strategic probabilities are in range of $[0,\ 1]$. This vector equation requires the following four probabilistic constrains:

$$\begin{aligned} p_1 &= 1 - \varphi(\chi - 1)(R - P), \\ p_2 &= 1 - \varphi[(P - S) + \chi(T - P)], \\ p_3 &= \varphi[\chi(P - S) + (T - P)], \\ p_4 &= 0. \end{aligned} \tag{4}$$

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

$$E_X - P = \chi(E_Y - P). \tag{5}$$

Both E_X and E_Y are maximized when Y fully cooperates (i.e., $\mathbf{p} = (1,1,1,1)$). Thus, player X's maximum payoff by using the χ-extortion strategy is

$$E_X = \frac{P(T-R) + \chi[R(T-S) - P(T-R)]}{(T-R) + \chi(R-S)}.$$

If the payoff vector $\mathbf{S}_X = (R,S,T,P)$ is arbitrary, there will be more than four parameters determining the extortion strategies as well as the payoff E_X. To simplify the analysis, Press and Dyson[10] fixed the PD payoff vector as the conventional $\mathbf{S}_X = (3,0,5,1)$, then they obtain the solution for extortion strategies as

$$\begin{aligned} p_1 &= 1 - 2\varphi(\chi - 1), \\ p_2 &= 1 - \varphi(4\chi + 1), \\ p_3 &= \varphi(4\chi + 1), \\ p_4 &= 0. \end{aligned} \tag{6}$$

Under this extortion strategy, the maximum payoff for player X is $E_X = (2+13\chi)/(2+3\chi)$, while the maximum payoff for player Y is $E_Y = (12+3\chi)/(2+3\chi)$. If player X adopts such extortion strategies, when Y adjusts her own strategy to increase the payoff, she will increase X's payoff even more, and when she achieves her own maximum payoff value by fully cooperating, X's payoff is also maximized. When player X adopts such an extortion strategy, player Y will be helpless even she knows cooperation benefits her opponent more than herself. This is the reason why such strategies are called extortion strategies.

4. Analysis

4.1. Robustness of extortion strategies

A player Y is called evolutionary if she is trying to improve her own payoff step by step by following certain optimization schemes. When the player X is confronting such an evolutionary player Y, whether an extortion strategy still brings her the maximum payoff? Although facing an extortion strategy player X, both players' payoffs will be maximized when Y fully cooperates, the evolutionary player Y may not realize this and only gradually change her strategy $\mathbf{q}$ to improve her payoff E_Y step by step. However, the evolutionary path of $\mathbf{q}$ is not unique and some paths may result in local peaks of E_Y, then the evolution of $\mathbf{q}$ is stopped, leaving both Y and X smaller payoffs than what the later expects. Therefore, confronting an evolutionary player Y, player X should make sure that her extortion strategy can afford positive gradients for Y's cooperative behaviors, such that Y would never

stop changing her strategy until the global maximums of both players' payoffs are reached. Otherwise, player X would continue monitoring Y's action in each stage game and choose a very complicated strategy to respond to Y's actions, other than simply adopting the extortion strategy.

Press and Dyson[10] conjectured that, if player X adopts an extortion strategy, the directional derivatives of E_Y with respect to $\mathbf{q}$ are positive everywhere, thus there always exist evolutionary paths that will lead to the maximum payoff for X. They numerically checked that the evolutionary paths never end up at a certain local optimum. Chen and Zinger[14] proved this conjecture in a very strong form. They concluded that, under player X's extortion strategy, for all kinds of PD payoff vectors, all kinds of initial game states and all kinds of strategies $\mathbf{q}$, every evolutionary path of Y will lead to the maximum payoffs of both players, even though Y may not end up with fully cooperation. That is to say, it is safe for X to take the extortion strategy as long as Y will not stop evolving her strategy, regardless of which evolutionary path Y will take. Their result even holds in some degenerated cases, where Press and Dyson's analysis does not apply, showing that the extortion strategies are extremely robust against evolutionary players.

4.2. Cooperative ZD strategies

The original ZD strategies are not necessarily promoting cooperations, since the Markov process does not surely converge to a joint state *CC*. Akin[13] proposed a new type of strategies called good strategies, under which, for player X, her opponent Y cannot simultaneously obtain a payoff R for herself while assign X a payoff smaller than R. If player X chooses a good strategy, the best response for Y is to cooperate, leading the game to the mutual cooperation state CC. A good strategy is more efficient than traditional Nash strategies, since Nash strategies do not surely promote cooperation. For instance, a mixed strategy $\mathbf{p} = a(1,0,0,1) + (1-a)(1,1,0,0)$ is a good strategy, since for any strategy $\mathbf{q}$, $E_X = E_Y = R$. In addition, a strategy $\mathbf{p} = a(1,0,0,0) + (1-a)(1,1,0,0)$ is also a good strategy, while the strategies $\mathbf{p} = (1,0,0,1)$, $\mathbf{p} = (1,1,0,0)$ and $\mathbf{p} = (1,0,0,0)$ are not good strategies themselves.

The good strategies and ZD strategies overlap at certain regions. Akin[13] showed that, a strategy $\mathbf{p}=(1,p_2,p_3,p_4)$ is a Nash strategy if and only if

$$\frac{T-R}{R-S}p_3 \leq 1-p_2, \quad \frac{T-R}{R-P}p_4 \leq 1-p_2, \tag{7}$$

and the strategy $\mathbf{p}$ is a good strategy if and only if both inequalities are strictly satisfied. Strategies in the overlapping regions of ZD strategies and good strategies are called generous ZD strategies, which not only lead to a unilateral control, but also promote the mutual cooperation.

4.3. Evolutionary stability of extortion strategies

Adami and Hintze[15] pointed out that the extortion strategy is evolutionarily unstable. In an evolutionary game with a well-mixed population, individuals play with each other and tend to learn successful strategies. At the beginning of such a game, an extortioner may obtain higher payoff. However, as more individuals switch to extortioners, extortioners' payoffs will decrease. This is because when extortioners meet one another, the payoff for each side will be the mutual punishment payoff P. Therefore, when extortioners dominate the population, their payoffs will approximate the value of P, and thus the extortion strategies are evolutionary unstable. Hilbe, Nowak, and Sigmund[16] showed that although extortion strategy is unstable under the natural selection, it promotes the emergence of cooperative strategies such as WSLS. It is known that WSLS will be beat by All-D. However, in a population composed of WSLS players, defectors and extortioners, at the early stages, extortioners can invade the defectors via neutral drift and selection, and later WSLS players will replace the extortioners and dominate the population. Therefore, the extortioners act as catalyst that increases the cooperation level. Furthermore, Stewart and Plotkin[18] studied the evolutionary performance of GZD strategies (i.e., $\tilde{\mathbf{p}} = \varphi[(\mathbf{S}_X - R\mathbf{1}) - \chi(\mathbf{S}_Y - R\mathbf{1})]$, $\chi > 0$), and found that GZD and WSLS strategies are particularly successful in a large population. Szolnoki and Perc[26] showed that the extortion strategies can exist in a structured population.

5. Extensions

5.1. ZD strategies in multi-player games

A natural question is whether the ZD strategies still exist in the generalized multi-player games. Pan *et al.*[27] constructed the mathematical framework of multi-player ZD strategies and accordingly analyzed the features by using the well-known iterated public goods game (IPGG). In IPGG, there are N players, and in each stage game, each player chooses whether or not to invest a token into a public pot. The tokens in the pot are multiplied by a factor r ($1 < r < N$), which produces the "public good" payoff. The public good payoff is then equally distributed among the N players regardless of their investment. It is found that the ZD strategies still exist in multi-player games and it is possible for a ZD strategy player to pin or to extort the total payoff of the other $N-1$ players. However, different from the observations in IPD, there exist some unreported constrains related to the number of players N and multiplication factor r, which determine the payoff value of the ZD player, as well as the feasible regions for pinning and extortion strategies.

For the pinning strategies in the IPGG, the pinned total payoff of other players is no less than the total investment of

them, which indicates that the pinning strategies should be realized by rewarding the others. Through further analysis of the probabilistic constrains of pinning strategies, it is found that a player can pin the others' total payoff only when the group size N does not exceed an upper bound, or when the multiplication factor r is not too large. Generally, pinning multiple players' total payoff in the IPGG is more difficult than pinning a single opponent's payoff in the two-player IPD. The authors also proved that, like in two-player games, a ZD strategy player also cannot unilaterally set her own payoff to a fixed value in the multi-player games.

For the extortion strategies, given any multiplication factor r, the lower bound of the extortion ratio χ is always 1. When $r > N/(N-1)$, χ also has its upper bound $r/[r(N-1)-N]$. By inspecting this upper bound, one can see that, when the multiplication factor is fixed, if there are more players involved in the game (i.e., N is larger), the feasible region of extortion will become smaller. This means the increase in the number of players will reduce the ability of an extortioner. Moreover, when the number of players N is fixed, a larger multiplication factor r will result in a smaller feasible region of extortion. That is to say, increasing the reward for players will also weaken the ability of extortioners.

Hilbe *et al.*[28] considered the ZD strategy alliances in multi-player games. They showed that, when some players constitute a strategic alliance and jointly launch a ZD strategy, a linear relationship will be enforced between the average payoff in the alliance and the payoffs of other players. They investigated the performance of four kinds of ZD strategy alliances, including the fair alliance, the extortionate alliance, the generous alliance, and the pinning alliance, and found that the effects of a ZD strategy alliance depend on the size of the alliance, the type of the social dilemma, and on the different strategies.

5.2. ZD strategies under noise

Rather than the perfect environment of Press and Dyson's original case, with a more realistic consideration, the game may contain noises which result in players' observation errors. Hao *et al.*[29] introduced the noise structure into the ZD strategy framework, and correspondingly proposed the generalized ZD strategies under observation error. They studied the Markov processes of the noisy repeated games and found that the determinants of the generalized transition matrixes still possess the Press–Dyson feature, although each column is involved in a more complicated structure. Based on this novel model, the authors formally derived and analyzed the pinning strategies and extortion strategies under noise. They found that, although affected by the observation error, it is still possible for a player to enforce these two specializations of ZD strategies to control its opponent's payoff.

For pinning strategies, both the pinned value of player Y's payoff and the feasible region of player X's strategy are limited by the noise distribution. In an environment with a weak noise, the corresponding pinning strategies generally exist, and player X's strategy as well as the noise distribution jointly determine player Y's payoff. However, when there is a very strong noise, the probability constrains for the pinning strategies are not essentially satisfied, and consequently, it may be impossible for X to set Y's payoff to be a fixed value. Furthermore, a pinning strategy under a specific noise may even fail to enforce a linear relationship between players' payoffs when the noise becomes relatively stronger. Intuitively, by using a same pinning strategy, as the noise level increases, the feasible region of the two players' payoff pairs expands gradually from a line to a plane.

The extortion strategies also exist in noisy environment. However, similar to that for the pinning strategies, the feasible region and the control function of extortion strategies are strongly constrained by the noise distribution. Recall that in the perfect environment without noise, an extortion line starts from the mutual punishment point (P,P) but never crosses the (P,P)–(R,R) segment (see Fig. 1(b)). However, with the noise introduced, the starting point of an extortion line will be changed due to the observation error. It should start at a point which has a small distance to (P,P), and the distance is totally determined by the noise distribution. As the noise level increases, the slop of the corresponding extortion line will become smaller, even to a negative value.

6. Challenges

Press and Dyson's discovery of ZD strategy have changed the researches on game theory. Realizing ZD strategies, a player will be powerful enough to exert her unilateral control on other players and dominate in the repeated games. The theory discovered by Press and Dyson applies not only to ZD strategies in games. Further research on Markov process and complex system could be potentially inspired by their contributions. We believe that the ZD strategy related research is still in its infancy, and we will continue to devote to this field. To the best of our knowledge, the immediate challenges include but are not limited to the following aspects.

ZD strategies in more general game models For example, how ZD strategies perform in other two-player and two-action games besides prisoner's dilemma? In two-player and two-action games, there are four parameters (R,S,T,P) characterizing the payoff structures. When these parameters change, the prisoner's dilemma may shift to other games. Thus, it is crucial to investigate whether the properties of ZD strategies still hold, and what are the new constrains when the payoff parameters vary. Furthermore, like in multi-player games, the ZD strategy in games with more than two actions

is another kind of generalization of Press and Dyson's framework. How to define the multi-action ZD strategies and what are the properties of them are open problems.

Extensions of fundamental ZD strategy frameworks The original ZD strategies are derived by satisfying several probabilistic constrains. Thus, finding ZD strategies is essentially a constrain satisfaction problem. Although some strategies may not satisfy the probability constrains, they may locate very close to the feasible region of ZD strategies. It is conjectured that, near the feasible region of ZD strategies, although a strategy cannot force a linear relationship between the players' payoffs, it can at least restrict the opponent's payoffs to a very small region, which seems to be a narrow ecliptic plane expanded from the ZD strategy line. Comparing with ZD strategies, this kind of strategies can realize a relatively weaker control of the opponent's payoff. Thus, we call them the Semi-ZD strategies. Furthermore, the condition of ZD strategies' existence relies on the convergency of the Markov process. When the game is not assumed to be infinitely repeated, the stationary distribution may not essentially exist. Thus, how to extend the original ideas into finitely repeated games is another open question.

Evolutionary dynamics of ZD strategies In the sense of evolutionary games, both the extortion strategies and the compromising strategies have already been studies. However, there are very few papers addressing how normal ZD strategies performs. Especially, how the evolutionary stability changes when normal ZD strategies are leaving the extortion side, but approaching the compromising side? This issue probably has relationship with the ZD strategy player's toughness. Furthermore, how different ZD strategies perform when they encounter with each other? How ZD strategies work in the networked evolutionary games? What are the effects of different network structures on the evolutionary dynamics of ZD strategies? All these issues may bring the evolutionary games with a new start.

References

[1] Axelrod R and Hamilton W D 1981 *Science* **211** 1390
[2] Axelrod R 1984 *The Evolution of Cooperation* (New York: Basic Book)
[3] Axelrod R and Dion D 1988 *Science* **242** 1385
[4] Nowak M A 2006 *Science* **314** 1560
[5] Szabó G and Fath G 2007 *Phys. Rep.* **446** 97
[6] Kendall G, Yao X and Chong S Y 2007 *The Iterative Prisoners' Dilemma: 20 Years On* (Singapore: World Scientific)
[7] Hardin G 1968 *Science* **162** 1243
[8] Mailath G and Samuelson L 2006 *Repeated Games and Reputations: Long-Run Relationships* (New York: Oxford University Press)
[9] Kandori M 1992 *Review of Economic Studies* **59** 63
[10] Press W H and Dyson F J 2012 *Proc. Acad. Natl. Sci. USA* **109** 10409
[11] Stewart A J and Plotkin J B 2012 *Proc. Acad. Natl. Sci. USA* **109** 10134
[12] Hayes B 2013 *American Scientist* **101** 422
[13] Akin E 2012 arXiv:1211.0969 [math.DS]
[14] Chen J and Zinger A (unpublished)
[15] Adami C and Hintze A 2013 *Nat. Commun.* **4** 3193
[16] Hilbe C, Nowak M A and Sigmund K 2013 *Proc. Acad. Natl. Sci. USA* **110** 6913
[17] Roemheld L 2013 arXiv:1308.2576 [cs.GT]
[18] Stewart A J and Plotkin J B 2013 *Proc. Acad. Natl. Sci. USA* **110** 15348
[19] Hilbe C, Nowak M A and Traulsen A 2013 *PLoS ONE* **8** e77886
[20] Neumann J V and Morgenstern O 1944 *Theory of Games and Economic Behavior* (Princeton: Princeton University Press)
[21] Nash J F 1950 *Proc. Acad. Natl. Sci. USA* **36** 48
[22] Nowak M A and Sigmund K 1992 *Nature* **355** 250
[23] Nowak M A and Sigmund K 1993 *Nature* **364** 56
[24] Smith J M 1982 *Evolution and the Theory of Games* (Cambridge: Cambridge University Press)
[25] Nowak M A, Sasaki A, Taylor C and Fudenberg D 2004 *Nature* **428** 646
[26] Szolnoki A and Perc M 2014 *Phys. Rev. E* **89** 022804
[27] Pan L, Hao D, Rong Z and Zhou T 2014 arXiv:1402.3542 [cs.GT]
[28] Hilbe C, Traulsen A, Wu B and Nowak M A 2014 arXiv:1404.2886 [q-bio.PE]
[29] Hao D, Rong Z, Pan L and Zhou T (unpublished)