Full Length Article

On the implementation of zero-determinant strategies in repeated games

Masahiko Ueda

Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Yamaguchi 753-8511, Japan

A R T I C L E   I N F O

A B S T R A C T

Zero-determinant strategies are a class of strategies in repeated games which unilaterally control payoffs. Zero-determinant strategies have attracted much attention in studies of social dilemma, particularly in the context of evolution of cooperation. So far, not only general properties of zero-determinant strategies have been investigated, but zero-determinant strategies have been applied to control in the fields of information and communications technology and analysis of imitation. Here, we further deepen our understanding on general mathematical properties of zero-determinant strategies. We first prove that zero-determinant strategies, if exist, can be implemented by some one-dimensional transition probability. Next, we prove that, if a two-player game has a non-trivial potential function, a zero-determinant strategy exists in its repeated version. These results assist us to implement zero-determinant strategies in broader situations.

1. Introduction

Zero-determinant (ZD) strategies are a class of strategies in repeated games which unilaterally enforce linear relations between payoffs. They were originally discovered in the infinitely repeated prisoner's dilemma games [1]. One of the examples is the equalizer strategy, which unilaterally controls the payoff of the opponent. Another example is the extortionate strategy, which cannot be beaten. There is also a ZD strategy called the generous ZD strategy, which cannot win but promotes mutual cooperation [2]. After the discovery of ZD strategies, properties of ZD strategies have extensively been investigated in the context of evolution of cooperation [3,4,2,5–8], including evolution on structured populations [9–11], together with human experiments of ZD strategies [12–14]. Furthermore, beyond the prisoner's dilemma game, ZD strategies were also discovered in multi-player social dilemma games [15,16] and a continuous donation game [17]. These results have suggested that ZD strategies can be used to control payoffs in a broader class of social dilemma situations.

At the same time as such investigation of ZD strategies in social dilemma situations, general properties of ZD strategies have also been studied. For example, for games with a discount factor, it has been recognized that the range of possible ZD strategies shrinks compared with no discounting case [18,17,19,20]. It was also found that effect of imperfect monitoring about the previous actions reduces the range of their existence, compared to perfect monitoring case [21,22]. Difference between these two effects in the prisoner's dilemma game was detailly investigated in Ref. [23]. Existence of ZD strategies in asynchronous-update social dilemma games was also shown in Ref. [24]. Furthermore, ZD strategies in general multi-player multi-action games were introduced in Ref. [25], and their mathematical properties in situations where several players use ZD strategies were also studied [22], which found that possible ZD strategies are constrained by consistency of linear payoff relations. A necessary and sufficient condition for

E-mail address: m.ueda@yamaguchi-u.ac.jp.

the existence of ZD strategies was recently specified [26]. Moreover, there are several attempts to extend the ability of payoff control of ZD strategies so as to control the moments of payoffs [27], time correlation functions of payoffs [28], and conditional expectations of payoffs [29].

Apart from these theoretical studies on general properties of ZD strategies, there are many applications of ZD strategies. In the field of information and communications technology, ZD strategies have been applied in resource sharing in wireless networks [30,31], smart grid systems [32], mining in blockchain [33,34], crowdsourcing [35,36], mobile crowdsensing [37], the Internet of Things [38], cloud computing [39], federated learning [40], and data trading [41], in order to unilaterally realize some favorable states. Another field of application is analysis of imitation in behavioral science [42,43]. Originally, in the prisoner's dilemma game, it was pointed out that the Tit-for-Tat (TFT) strategy [44,45], which imitates the previous action of the opponent, is a ZD strategy which unilaterally equalizes the payoffs of two players [1]. Because it has been known that TFT is unbeatable in two-player symmetric potential games [46], relation between unbeatability and ZD strategies was focused on, and it was indeed shown that unbeatable TFT is a ZD strategy [42]. Furthermore, relation between the existence of unbeatable imitation [47] and that of unbeatable ZD strategies has also been shown in two-player symmetric games [26] and in multi-player totally symmetric games [43]. These application studies show that the concept of ZD strategies can be useful as methods to control multi-agent systems and to analyze human behavior.

Although the concept of ZD strategy has been widely used, its general mathematical properties have not yet been sufficiently clear. Clarifying its properties will enable us to implement ZD strategies in broader situations which were not in the scope of application of ZD strategies [48,49]. In this paper, we further clarify mathematical properties of ZD strategies. Particularly, we first show that ZD strategies, if exist, can be realized by one-dimensional transition probability. This result enables us to implement ZD strategies even if an action set is restricted to one-dimension space. Second, we prove that ZD strategies controlling payoffs exist in two-player non-trivial potential games. This result can be regarded as an extension of the result in two-player symmetric potential games [42] to asymmetric cases.

This paper is organized as follows. In Section 2, we introduce basic concepts such as repeated games, zero-determinant strategies, and potential games. In Section 3, we prove our first theorem about implementation of ZD strategies by one-dimensional transition probability. In Section 4, we provide our second theorem about the existence of ZD strategies controlling payoffs in two-player non-trivial potential games. This result is also demonstrated by using an asymmetric prisoner's dilemma game in this section. In Section 5, we discuss relations between our results and previous results. Section 6 is devoted to concluding remarks.

2. Preliminaries

2.1. Repeated games

A game in strategic form is defined as $G := \left(\mathcal{N}, \{A_j\}_{j \in \mathcal{N}}, \{s_j\}_{j \in \mathcal{N}} \right)$, where $\mathcal{N}$ is the set of players, A_j is the set of actions of player $j \in \mathcal{N}$, and $s_j : \prod_{j \in \mathcal{N}} A_j \to \mathbb{R}$ is the payoff function of player $j \in \mathcal{N}$ [50,51]. We consider the situation that $\mathcal{N} := \{1, \cdots, N\}$ and $A_j := \left\{ b_j^{(1)}, \cdots, b_j^{(L_j)} \right\}$ for all $j \in \mathcal{N}$, where $L_j \in \mathbb{N}$ is the number of action of player j. For simplicity, we write $\mathcal{A} := \prod_{j=1}^{N} A_j$ and $\boldsymbol{a} := (a_1, \cdots, a_N) \in \mathcal{A}$, and call $\boldsymbol{a}$ an action profile. Below we also use notation $a_{-j} := (a_1, \cdots, a_{j-1}, a_{j+1}, \cdots, a_N)$ and $A_{-j} := \prod_{k \neq j} A_k$. When we want to emphasize a_j in $\boldsymbol{a}$, we write $\boldsymbol{a} = (a_j, a_{-j})$.

When G is infinitely repeated and players choose actions at each round according to all histories of actions, such game is called an (infinitely) repeated game. We write an action of player j at round $t \geq 1$ as $a_j^{(t)}$. Below Δ_L denotes a probability L-simplex. As strategies in repeated games, we consider behavior strategies $\mathcal{T}_j := \left\{ T_j^{(t)} \right\}_{t=1}^{\infty}$ $(\forall j \in \mathcal{N})$, where $T_j^{(t)} : \mathcal{A}^{t-1} \to \Delta_{L_j}$ with $t \geq 2$ is the conditional probability at t-th round and $T_j^{(1)} \in \Delta_{L_j}$ is the initial probability. Similarly to the original work [1], in this paper, we only consider the case that payoffs in the infinitely repeated game are given by the time-averaged payoffs (that is, no discounting)

$$S_j := \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[s_j\left(\boldsymbol{a}^{(t)}\right) \right] \quad (\forall j \in \mathcal{N}), \tag{1}$$

where $\mathbb{E}[\cdots]$ is the expectation of the quantity $\cdots$ with respect to strategies of all players.

We write the joint probability distribution of action profiles from first round to t-th round as $P\left(\boldsymbol{a}^{(t)}, \cdots, \boldsymbol{a}^{(1)}\right)$. We also define the marginal probability distribution of an action profile at t-th round obtained from the joint probability distribution by

$$P_t\left(\boldsymbol{a}^{(t)}\right) := \sum_{\boldsymbol{a}^{(t-1)}} \cdots \sum_{\boldsymbol{a}^{(1)}} P\left(\boldsymbol{a}^{(t)}, \cdots, \boldsymbol{a}^{(1)}\right) \quad \left(\forall \boldsymbol{a}^{(t)} \in \mathcal{A}\right). \tag{2}$$

If the limit probability distribution

$$P^*(\boldsymbol{a}) := \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} P_t(\boldsymbol{a}) \quad (\forall \boldsymbol{a} \in \mathcal{A}), \tag{3}$$

exists, we write the expectation of the quantity $D : \mathcal{A} \to \mathbb{R}$ with respect to P^* by $\langle D \rangle^*$. For such case, the payoffs (1) are described as $S_k = \langle s_k \rangle^*$ for all $k \in \mathcal{N}$.

2.2. Zero-determinant strategies

Here, we define a *time-independent memory-one strategy* of player $j \in \mathcal{N}$ as a strategy such that

$$T_j^{(t)}\left(a_j^{(t)}|\boldsymbol{a}^{(t-1)}, \cdots, \boldsymbol{a}^{(1)}\right) = T_j\left(a_j^{(t)}|\boldsymbol{a}^{(t-1)}\right) \quad (\forall t \geq 2) \tag{4}$$

for all $a_j^{(t)}$, $\boldsymbol{a}^{(t-1)}$, $\cdots$, $\boldsymbol{a}^{(1)}$, where $T_j : \mathcal{A} \to \Delta_{L_j}$. That is, a player using time-independent memory-one strategy chooses an action in each round only by using information of a previous action profile. Furthermore, for time-independent memory-one strategies T_j of player j, we introduce the *Press-Dyson vectors* [6,17,22]

$$\hat{T}_j\left(a_j|\boldsymbol{a}'\right) := T_j\left(a_j|\boldsymbol{a}'\right) - \delta_{a_j, a_j'} \quad \left(\forall a_j \in A_j, \forall \boldsymbol{a}' \in \mathcal{A}\right), \tag{5}$$

where $\delta_{a,a'}$ is the Kronecker delta. Note that the Press-Dyson vectors describe the difference between the time-independent memory-one strategy T_j and the strategy which repeats one's previous action (the Repeat strategy [6]) $\delta_{a_j, a_j'}$.

We remark that, because T_j is probability, Press-Dyson vectors generally satisfy

$$\sum_{a_j} \hat{T}_j\left(a_j|\boldsymbol{a}'\right) = 0 \quad (\forall \boldsymbol{a}') \tag{6}$$

and

$$\begin{cases} -1 \leq \hat{T}_j\left(a_j|\boldsymbol{a}'\right) \leq 0 & \left(a_j = a_j'\right) \\ 0 \leq \hat{T}_j\left(a_j|\boldsymbol{a}'\right) \leq 1 & \left(a_j \neq a_j'\right). \end{cases} \tag{7}$$

Now we introduce zero-determinant strategies in repeated games [1,17,29].

Definition 1. A time-independent memory-one strategy of player j is an (extended) *zero-determinant (ZD) strategy* controlling the quantity $B : \mathcal{A} \to \mathbb{R}$ when its Press-Dyson vectors can be written in the form

$$\sum_{a_j} c_{a_j} \hat{T}_j\left(a_j|\boldsymbol{a}'\right) = B\left(\boldsymbol{a}'\right) \quad (\forall \boldsymbol{a}' \in \mathcal{A}) \tag{8}$$

with some nontrivial coefficients $\left\{c_{a_j}\right\}$ (that is, not $c_{b_1} = \cdots = c_{b_{L_j}} = \text{const.}$) and B is not identically zero.

Counterintuitively, the ZD strategy unilaterally enforces

$$\langle B \rangle^* = 0, \tag{9}$$

that is, $\langle B \rangle^* = 0$ holds independently of the strategies of other players [1,17,29]. Originally, ZD strategies controlling $B(\boldsymbol{a}) = \sum_{j=1}^{N} \alpha_j s_j(\boldsymbol{a}) + \alpha_0$ with some coefficients $\left\{\alpha_j\right\}_{j=0}^{N}$ have been investigated.

An existence condition of ZD strategies is given by the following proposition:

Proposition 1 (*[26]*). *A ZD strategy of player j controlling B exists if and only if there exist two different actions $\overline{a}_j, \underline{a}_j \in A_j$ of player j such that*

$$B\left(\overline{a}_j, a_{-j}\right) \geq 0 \quad (\forall a_{-j})$$
$$B\left(\underline{a}_j, a_{-j}\right) \leq 0 \quad (\forall a_{-j}), \tag{10}$$

and B is not identically zero.

2.3. Potential games

The concept of potential games was originally introduced in Ref. [52] inspired by physics.

Definition 2. A game G is an *(exact) potential game* when there exists a common function $\Phi(\boldsymbol{a})$ satisfying

$$s_j(a_j, a_{-j}) - s_j(a_j', a_{-j}) = \Phi(a_j, a_{-j}) - \Phi(a_j', a_{-j}) \quad (\forall a_j, \forall a_j', \forall a_{-j}) \tag{11}$$

for all player j.

The function Φ is called a potential function. It should be noted that the Nash equilibrium $\boldsymbol{a}^*$ is defined by the condition

$$s_j\left(a_j^*, a_{-j}^*\right) \geq s_j\left(a_j, a_{-j}^*\right) \quad (\forall j, \forall a_j). \tag{12}$$

For a potential game, since this condition is rewritten as

$$\Phi\left(a_j^*, a_{-j}^*\right) \geq \Phi\left(a_j, a_{-j}^*\right) \quad (\forall j, \forall a_j), \tag{13}$$

Nash equilibria are realized as local maxima of a potential function. A sequence of actions improving the payoffs can be regarded as that increasing the value of a potential function.

For subsets $\forall \mathcal{N}' \subseteq \mathcal{N}$, we write $a_{\mathcal{N}'} := \left(a_j\right)_{j \in \mathcal{N}'}$ and $A_{\mathcal{N}'} := \prod_{j \in \mathcal{N}'} A_j$. Ui proved that a potential function can be described as the sum of interaction potentials.

Proposition 2 ([53]). *G is a potential game if and only if there exist interaction potentials $\left\{\phi_{\mathcal{N}'} | \phi_{\mathcal{N}'} : A_{\mathcal{N}'} \to \mathbb{R}, \mathcal{N}' \subseteq \mathcal{N}\right\}$ such that*

$$s_j(\boldsymbol{a}) = \sum_{\mathcal{N}' \subseteq \mathcal{N}}^{j \in \mathcal{N}'} \phi_{\mathcal{N}'}\left(a_{\mathcal{N}'}\right) \tag{14}$$

for all $j \in \mathcal{N}$ and all $\boldsymbol{a} \in \mathcal{A}$. A potential function is given by

$$\Phi(\boldsymbol{a}) = \sum_{\mathcal{N}' \subseteq \mathcal{N}} \phi_{\mathcal{N}'}\left(a_{\mathcal{N}'}\right). \tag{15}$$

3. Realization of ZD strategies by one-dimensional transition probability

3.1. Main result

First, we prove that, if ZD strategies exist, they can be implemented by one-dimensional transition probability. Here, we call transition probability of a_j *one-dimensional* when transition of a_j occurs only between nearest neighbor sites in a_j-coordinate. We consider the situation that a ZD strategy of player j controlling B exists. We can relabel the name of actions of player j from $\left\{b_j^{(1)}, \cdots, b_j^{(L_j)}\right\}$ to $\left\{1, \cdots, L_j\right\}$ so as to satisfy $\overline{a}_j = 1$ and $\underline{a}_j = L_j$ in Eq. (10). (The rest of actions is renamed arbitrarily.) Below we redefine $A_j = \left\{1, \cdots, L_j\right\}$.

Theorem 1. *If a ZD strategy of player j controlling B exists, it can be realized by one-dimensional transition probability.*

Proof. According to Proposition 1, a necessary and sufficient condition for the existence of a ZD strategy is given by Eq. (10), that is,

$$B\left(1, a_{-j}\right) \geq 0 \quad \left(\forall a_{-j}\right)$$
$$B\left(L_j, a_{-j}\right) \leq 0 \quad \left(\forall a_{-j}\right) \tag{16}$$

and B is not identically zero. Below we write the total number of action profiles as $M := \prod_{k=1}^{N} L_k$. For simplicity, we introduce a vector notation of the quantity $D : \mathcal{A} \to \mathbb{R}$ by $\boldsymbol{D} := (D(\boldsymbol{a}))_{\boldsymbol{a} \in \mathcal{A}} \in \mathbb{R}^M$. It should be noted that a trivial identity

$$D(\boldsymbol{a}') = \sum_{a_j \in A_j} \sum_{d=\pm 1} D(\boldsymbol{a}') \mathbb{I}(d\,D(\boldsymbol{a}') > 0) \mathbb{I}(a_j' = a_j) \tag{17}$$

holds for all $\boldsymbol{a}' \in \mathcal{A}$, where $\mathbb{I}(\cdots)$ is an indicator function which returns 1 if $\cdots$ holds, and 0 otherwise. By using the above vector notation, this identity is rewritten as

$$\boldsymbol{D} = \sum_{a_j} \sum_{d=\pm 1} [\boldsymbol{D}]_{a_j, \mathrm{sgn}(d)}, \tag{18}$$

where we have defined

$$[\boldsymbol{D}]_{a_j, \mathrm{sgn}(d)} := \left(D(\boldsymbol{a}') \mathbb{I}(d\,D(\boldsymbol{a}') > 0) \mathbb{I}(a_j' = a_j)\right)_{\boldsymbol{a}' \in \mathcal{A}}$$
$$\left(a_j \in A_j, d \in \{1, -1\}\right). \tag{19}$$

For the quantity $\boldsymbol{B} \in \mathbb{R}^M$, our assumption (16) implies

$$\boldsymbol{B} = \sum_{a_j \neq 1, L_j} \sum_{d=\pm 1} [\boldsymbol{B}]_{a_j, \mathrm{sgn}(d)} + [\boldsymbol{B}]_{L_j, -} + [\boldsymbol{B}]_{1, +}. \tag{20}$$

We also collectively write the Press-Dyson vectors of player j by $\hat{\boldsymbol{T}}_j\left(a_j\right) := \left(\hat{T}_j\left(a_j|\boldsymbol{a}'\right)\right)_{\boldsymbol{a}'\in\mathcal{A}}$ for all $a_j \in A_j$.

We remark that the quantity

$$W := \max_{\boldsymbol{a}\in\mathcal{A}} |B(\boldsymbol{a})| \tag{21}$$

satisfies $W \neq 0$ because of our assumption that B is not identically zero. We now consider the following strategy of player j:

$$\begin{aligned}
\hat{\boldsymbol{T}}_j(1) &= \frac{1}{W}\left(-[\boldsymbol{B}]_{1,+} - [\boldsymbol{B}]_{2,-}\right) \\
\hat{\boldsymbol{T}}_j\left(a_j\right) &= \frac{1}{W}\left(-[\boldsymbol{B}]_{a_j,+} + [\boldsymbol{B}]_{a_j,-} - [\boldsymbol{B}]_{a_j+1,-} + [\boldsymbol{B}]_{a_j-1,+}\right) \\
&\quad \left(2 \leq a_j \leq L_j - 1\right) \\
\hat{\boldsymbol{T}}_j\left(L_j\right) &= \frac{1}{W}\left([\boldsymbol{B}]_{L_j,-} + [\boldsymbol{B}]_{L_j-1,+}\right).
\end{aligned} \tag{22}$$

It should be noted that this strategy is different from a strategy originally introduced in Ref. [26] for the proof of Proposition 1, which is not one-dimensional. We check that Eq. (22) is indeed Press-Dyson vectors of some strategy. By the definitions (19) and (21), it satisfies Eq. (7). Furthermore, we obtain

$$\begin{aligned}
\sum_{a_j} \hat{\boldsymbol{T}}_j\left(a_j\right) &= \frac{1}{W}\left(-\sum_{a_j=1}^{L_j-1}[\boldsymbol{B}]_{a_j,+} + \sum_{a_j=2}^{L_j}[\boldsymbol{B}]_{a_j,-}\right. \\
&\quad \left.-\sum_{a_j=1}^{L_j-1}[\boldsymbol{B}]_{a_j+1,-} + \sum_{a_j=2}^{L_j}[\boldsymbol{B}]_{a_j-1,+}\right) \\
&= \boldsymbol{0},
\end{aligned} \tag{23}$$

which implies Eq. (6).

Next, we show that Eq. (22) is a ZD strategy. By calculating

$$\begin{aligned}
\sum_{a_j} a_j \hat{\boldsymbol{T}}_j\left(a_j\right) &= \frac{1}{W}\left(-\sum_{a_j=1}^{L_j-1} a_j [\boldsymbol{B}]_{a_j,+} + \sum_{a_j=2}^{L_j} a_j [\boldsymbol{B}]_{a_j,-}\right. \\
&\quad \left.-\sum_{a_j=1}^{L_j-1} a_j [\boldsymbol{B}]_{a_j+1,-} + \sum_{a_j=2}^{L_j} a_j [\boldsymbol{B}]_{a_j-1,+}\right) \\
&= \frac{1}{W}\left(\sum_{a_j=1}^{L_j-1} [\boldsymbol{B}]_{a_j,+} + \sum_{a_j=2}^{L_j} [\boldsymbol{B}]_{a_j,-}\right) \\
&= \frac{1}{W}\boldsymbol{B},
\end{aligned} \tag{24}$$

we can check that the strategy is indeed a ZD strategy (Definition 1).

Finally, we check that Eq. (22) is realized by one-dimensional transition probability. It is trivial from the definition (19) and the definition of the Press-Dyson vectors. Explicitly, each component of the transition probability of the ZD strategy (22) is

$$\begin{aligned}
T_j\left(1|\boldsymbol{a}'\right) &= \delta_{1,a_j'} + \frac{1}{W}\left(-\left|B\left(\boldsymbol{a}'\right)\right|\delta_{1,a_j'} + \left|B\left(\boldsymbol{a}'\right)\right|\mathbb{I}(B(\boldsymbol{a}')<0)\delta_{2,a_j'}\right) \\
T_j\left(a_j|\boldsymbol{a}'\right) &= \delta_{a_j,a_j'} + \frac{1}{W}\left(-\left|B\left(\boldsymbol{a}'\right)\right|\delta_{a_j,a_j'} + \left|B\left(\boldsymbol{a}'\right)\right|\mathbb{I}(B(\boldsymbol{a}')<0)\delta_{a_j+1,a_j'}\right. \\
&\quad \left.+\left|B\left(\boldsymbol{a}'\right)\right|\mathbb{I}(B(\boldsymbol{a}')>0)\delta_{a_j-1,a_j'}\right) \\
&\quad \left(2 \leq a_j \leq L_j - 1\right) \\
T_j\left(L_j|\boldsymbol{a}'\right) &= \delta_{L_j,a_j'} + \frac{1}{W}\left(-\left|B\left(\boldsymbol{a}'\right)\right|\delta_{L_j,a_j'} + \left|B\left(\boldsymbol{a}'\right)\right|\mathbb{I}(B(\boldsymbol{a}')>0)\delta_{L_j-1,a_j'}\right),
\end{aligned} \tag{25}$$

which is one-dimensional. $\square$

We remark that $\left\{C\hat{\boldsymbol{T}}_j(a_j)\right\}_{a_j\in A_j}$ with $0 < C \leq 1$, where $\left\{\hat{\boldsymbol{T}}_j(a_j)\right\}_{a_j\in A_j}$ is given by Eq. (22), is also a ZD strategy, due to Definition 1. In addition, the ZD strategy (25) increases a_j when $B(\boldsymbol{a}') > 0$ in the previous round and decreases a_j when $B(\boldsymbol{a}') < 0$ in the previous round under the boundary condition (16). This fact, which is similar to feedback control, provides an interpretation why a ZD strategy can unilaterally enforce Eq. (9). We also remark that each element of the transition probability (25) depends only on the quantity at the previous state $\boldsymbol{a}'$, like the trap model [54].

3.2. Continuum limit

Although informal, here we consider the continuum limit of this one-dimensional stochastic process. When all players use time-independent memory-one strategies, the time evolution is described by the Markov chain

$$P_{t+1}(\boldsymbol{a}) = \sum_{\boldsymbol{a}'} T(\boldsymbol{a}|\boldsymbol{a}') P_t(\boldsymbol{a}') \tag{26}$$

with the total transition probability

$$T(\boldsymbol{a}|\boldsymbol{a}') := \prod_{j=1}^{N} T_j(a_j|\boldsymbol{a}'). \tag{27}$$

We first consider the continuous time limit and then the continuous space limit. For this purpose, we assume that transition in T_j occurs with probability of order Δt for all j:

$$T_j(a_j|\boldsymbol{a}') = \delta_{a_j,a_j'} + \Delta t R_j(a_j|\boldsymbol{a}'), \tag{28}$$

where R_j is the transition rate. It should be noted that this assumption provides the meaning of the Press-Dyson vectors (5):

$$\hat{T}_j(a_j|\boldsymbol{a}') = \Delta t R_j(a_j|\boldsymbol{a}'). \tag{29}$$

According to Definition 1, as far as ZD strategies are concerned, multiplying a constant to the Press-Dyson vectors does not change properties of ZD strategies. Therefore, we can identify $\hat{T}_j$ and R_j.

By taking $\Delta t \to 0$ in Eq. (26), we obtain

$$\frac{\partial P}{\partial t}(\boldsymbol{a};t) = \sum_{j=1}^{N} \sum_{a_j'} R_j(a_j|a_j',a_{-j}) P(a_j',a_{-j};t), \tag{30}$$

where $P(\boldsymbol{a};t)$ is the probability distribution of $\boldsymbol{a}$ at real time $t \in \mathbb{R}$. We remark that the transition rate satisfies

$$\sum_{a_j} R_j(a_j|\boldsymbol{a}') = 0 \quad (\forall j, \forall \boldsymbol{a}'), \tag{31}$$

due to the normalization condition of Eq. (28).

Next, we assume that player j uses a ZD strategy, and consider the continuous space limit of Eq. (30) with respect to a_j. When R_j is given by the ZD strategy (22) with $1/W$ replaced by $\rho > 0$, the j-th term in the right-hand side of Eq. (30) is explicitly written as

$$\sum_{a_j'} R_j(a_j|a_j',a_{-j}) P(a_j',a_{-j};t)$$

$$= \rho \sum_{a_j'} \left[-\left|B(a_j',a_{-j})\right| \delta_{a_j,a_j'} + \left|B(a_j',a_{-j})\right| \mathbb{I}\left(B(a_j',a_{-j}) < 0\right) \delta_{a_j+1,a_j'} \right.$$

$$\left. + \left|B(a_j',a_{-j})\right| \mathbb{I}\left(B(a_j',a_{-j}) > 0\right) \delta_{a_j-1,a_j'} \right] P(a_j',a_{-j};t)$$

$$= \rho \left[-\left|B(a_j,a_{-j})\right| P(a_j,a_{-j};t) \right.$$

$$+ \left|B(a_j+1,a_{-j})\right| \mathbb{I}(B(a_j+1,a_{-j}) < 0) P(a_j+1,a_{-j};t)$$

$$\left. + \left|B(a_j-1,a_{-j})\right| \mathbb{I}(B(a_j-1,a_{-j}) > 0) P(a_j-1,a_{-j};t) \right] \tag{32}$$

for $2 \le a_j \le L_j - 1$. Here, we regard the length between two nearest neighbor sites in a_j-coordinate as Δa, and define $z_j := a_j \Delta a$. Furthermore, we redefine the probability distribution of (z_j, a_{-j}) by $\tilde{P}(z_j, a_{-j};t) := P(a_j, a_{-j};t)/\Delta a$, and the quantity B by $\tilde{B}(z_j, a_{-j}) := B(a_j, a_{-j})$. We assume that the function $\tilde{B}$ is smooth with respect to z_j in the limit $\Delta a \to 0$ with $\tilde{L}_j := L_j \Delta a$ and $\tilde{\rho} := \rho \Delta a$ fixed. Then, Eq. (32) is rewritten as

$$\sum_{a_j'} R_j(a_j|a_j',a_{-j}) \tilde{P}(a_j'\Delta a, a_{-j};t)$$

$$= \rho \left[-\left|\tilde{B}(z_j,a_{-j})\right| \tilde{P}(z_j,a_{-j};t) \right.$$

$$+ \left|\tilde{B}(z_j+\Delta a,a_{-j})\right| \mathbb{I}(\tilde{B}(z_j+\Delta a,a_{-j}) < 0) \tilde{P}(z_j+\Delta a,a_{-j};t)$$

$$\left. + \left|\tilde{B}(z_j-\Delta a,a_{-j})\right| \mathbb{I}(\tilde{B}(z_j-\Delta a,a_{-j}) > 0) \tilde{P}(z_j-\Delta a,a_{-j};t) \right]$$

$$= \rho \left[-\Delta a \frac{\partial}{\partial z_j} \left(\tilde{B} \left(z_j, a_{-j} \right) \tilde{P} \left(z_j, a_{-j}; t \right) \right) \right.$$

$$\left. + \frac{\Delta a^2}{2} \frac{\partial^2}{\partial z_j^2} \left(\left| \tilde{B} \left(z_j, a_{-j} \right) \right| \tilde{P} \left(z_j, a_{-j}; t \right) \right) + \mathcal{O}(\Delta a^3) \right] \tag{33}$$

$$\to -\tilde{\rho} \frac{\partial}{\partial z_j} \left(\tilde{B} \left(z_j, a_{-j} \right) \tilde{P} \left(z_j, a_{-j}; t \right) \right) \quad (\Delta a \to 0). \tag{34}$$

Therefore, in the limit $\Delta a \to 0$, we find that the dynamics of z_j is described by the time evolution equation:

$$\dot{z}_j = \tilde{\rho} \tilde{B} \left(z_j, a_{-j} \right). \tag{35}$$

Note that, when we calculate $\lim_{\tau \to \infty} 1/\tau \int_0^\tau dt$ of both sides, this equation becomes

$$0 = \tilde{\rho} \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau dt \, \tilde{B} \left(z_j(t), a_{-j}(t) \right) \tag{36}$$

because $0 \leq z_j \leq \tilde{L}_j$. This is why our ZD strategy unilaterally enforces $\langle B \rangle^* = 0$ irrespective of strategies of other players.

4. Existence of ZD strategies in two-player potential games

4.1. Main result

Next, we prove that ZD strategies controlling payoffs exist in two-player non-trivial potential games. According to Proposition 2, for two-player potential games, the payoffs and a potential is explicitly described as

$$s_1 (\boldsymbol{a}) = \phi_{\{1\}} \left(a_1 \right) + \phi_{\{1,2\}} \left(a_1, a_2 \right)$$
$$s_2 (\boldsymbol{a}) = \phi_{\{2\}} \left(a_2 \right) + \phi_{\{1,2\}} \left(a_1, a_2 \right) \tag{37}$$

and

$$\Phi (\boldsymbol{a}) = \phi_\emptyset + \phi_{\{1\}} \left(a_1 \right) + \phi_{\{2\}} \left(a_2 \right) + \phi_{\{1,2\}} \left(a_1, a_2 \right). \tag{38}$$

Theorem 2. *For every two-player potential game with $\phi_{\{1\}} \neq$ const. or $\phi_{\{2\}} \neq$ const., there exists a ZD strategy controlling*

$$s_1 (\boldsymbol{a}) - s_2 (\boldsymbol{a}) - \frac{\overline{\phi}_{\{1\}} + \underline{\phi}_{\{1\}}}{2} + \frac{\overline{\phi}_{\{2\}} + \underline{\phi}_{\{2\}}}{2}, \tag{39}$$

where we have defined

$$\overline{\phi}_{\{j\}} := \max_{a_j} \phi_{\{j\}} \left(a_j \right) \quad (\forall j)$$

$$\underline{\phi}_{\{j\}} := \min_{a_j} \phi_{\{j\}} \left(a_j \right) \quad (\forall j). \tag{40}$$

Proof. We first define

$$\overline{a}_j := \arg \max_{a_j} \phi_{\{j\}} \left(a_j \right) \quad (\forall j)$$

$$\underline{a}_j := \arg \min_{a_j} \phi_{\{j\}} \left(a_j \right) \quad (\forall j), \tag{41}$$

where ties may be broken arbitrarily. We consider the following two cases separately.

(i) $\overline{\phi}_{\{1\}} - \underline{\phi}_{\{1\}} \geq \overline{\phi}_{\{2\}} - \underline{\phi}_{\{2\}}$

We consider the quantity

$$B (\boldsymbol{a}) := s_1 (\boldsymbol{a}) - s_2 (\boldsymbol{a}) - \frac{\overline{\phi}_{\{1\}} + \underline{\phi}_{\{1\}}}{2} + \frac{\overline{\phi}_{\{2\}} + \underline{\phi}_{\{2\}}}{2}. \tag{42}$$

By using the property of potential games (37), this quantity is rewritten as

$$B (\boldsymbol{a}) = \phi_{\{1\}} \left(a_1 \right) - \phi_{\{2\}} \left(a_2 \right) - \frac{\overline{\phi}_{\{1\}} + \underline{\phi}_{\{1\}}}{2} + \frac{\overline{\phi}_{\{2\}} + \underline{\phi}_{\{2\}}}{2}. \tag{43}$$

Then we find that

$$B\left(\overline{a}_1, a_2\right) = \frac{\overline{\phi}_{\{1\}} - \underline{\phi}_{\{1\}}}{2} - \phi_{\{2\}}\left(a_2\right) + \frac{\overline{\phi}_{\{2\}} + \underline{\phi}_{\{2\}}}{2} \tag{44}$$

$$B\left(\underline{a}_1, a_2\right) = -\frac{\overline{\phi}_{\{1\}} - \underline{\phi}_{\{1\}}}{2} - \phi_{\{2\}}\left(a_2\right) + \frac{\overline{\phi}_{\{2\}} + \underline{\phi}_{\{2\}}}{2}. \tag{45}$$

By the definition of the maximum and the minimum, we obtain

$$B\left(\overline{a}_1, a_2\right) \geq B\left(\overline{a}_1, \overline{a}_2\right)$$
$$= \frac{\overline{\phi}_{\{1\}} - \underline{\phi}_{\{1\}}}{2} - \frac{\overline{\phi}_{\{2\}} - \underline{\phi}_{\{2\}}}{2}$$
$$\geq 0 \tag{46}$$

$$B\left(\underline{a}_1, a_2\right) \leq B\left(\underline{a}_1, \underline{a}_2\right)$$
$$= -\frac{\overline{\phi}_{\{1\}} - \underline{\phi}_{\{1\}}}{2} + \frac{\overline{\phi}_{\{2\}} - \underline{\phi}_{\{2\}}}{2}$$
$$\leq 0 \tag{47}$$

for all a_2. Furthermore, $B(\boldsymbol{a})$ is identically zero only if $\phi_{\{1\}}$ and $\phi_{\{2\}}$ are constant functions. Such situations are excluded by the assumption of theorem. Therefore, the two actions $\overline{a}_1$ and $\underline{a}_1$ can be regarded as those in Proposition 1, and a ZD strategy of player 1 exists.

(ii) $\overline{\phi}_{\{1\}} - \underline{\phi}_{\{1\}} \leq \overline{\phi}_{\{2\}} - \underline{\phi}_{\{2\}}$
We consider the quantity

$$B(\boldsymbol{a}) := s_2(\boldsymbol{a}) - s_1(\boldsymbol{a}) - \frac{\overline{\phi}_{\{2\}} + \underline{\phi}_{\{2\}}}{2} + \frac{\overline{\phi}_{\{1\}} + \underline{\phi}_{\{1\}}}{2}. \tag{48}$$

By using the property of potential games (37), this quantity is rewritten as

$$B(\boldsymbol{a}) = \phi_{\{2\}}\left(a_2\right) - \phi_{\{1\}}\left(a_1\right) - \frac{\overline{\phi}_{\{2\}} + \underline{\phi}_{\{2\}}}{2} + \frac{\overline{\phi}_{\{1\}} + \underline{\phi}_{\{1\}}}{2}. \tag{49}$$

Then we find that

$$B\left(a_1, \overline{a}_2\right) = \frac{\overline{\phi}_{\{2\}} - \underline{\phi}_{\{2\}}}{2} - \phi_{\{1\}}\left(a_1\right) + \frac{\overline{\phi}_{\{1\}} + \underline{\phi}_{\{1\}}}{2} \tag{50}$$

$$B\left(a_1, \underline{a}_2\right) = -\frac{\overline{\phi}_{\{2\}} - \underline{\phi}_{\{2\}}}{2} - \phi_{\{1\}}\left(a_1\right) + \frac{\overline{\phi}_{\{1\}} + \underline{\phi}_{\{1\}}}{2}. \tag{51}$$

By the definition of the maximum and the minimum, we obtain

$$B\left(a_1, \overline{a}_2\right) \geq B\left(\overline{a}_1, \overline{a}_2\right)$$
$$= \frac{\overline{\phi}_{\{2\}} - \underline{\phi}_{\{2\}}}{2} - \frac{\overline{\phi}_{\{1\}} - \underline{\phi}_{\{1\}}}{2}$$
$$\geq 0 \tag{52}$$

$$B\left(a_1, \underline{a}_2\right) \leq B\left(\underline{a}_1, \underline{a}_2\right)$$
$$= -\frac{\overline{\phi}_{\{2\}} - \underline{\phi}_{\{2\}}}{2} + \frac{\overline{\phi}_{\{1\}} - \underline{\phi}_{\{1\}}}{2}$$
$$\leq 0 \tag{53}$$

for all a_1. Furthermore, $B(\boldsymbol{a})$ is identically zero only if $\phi_{\{1\}}$ and $\phi_{\{2\}}$ are constant functions. Such situations are excluded by the assumption of theorem. Therefore, the two actions $\overline{a}_2$ and $\underline{a}_2$ can be regarded as those in Proposition 1, and a ZD strategy of player 2 exists.

Thus, we conclude that at least one ZD strategy exists under the assumption. We note that a linear relation enforced by the ZD strategy is

$$\langle s_1 \rangle^* - \frac{\overline{\phi}_{\{1\}} + \underline{\phi}_{\{1\}}}{2} = \langle s_2 \rangle^* - \frac{\overline{\phi}_{\{2\}} + \underline{\phi}_{\{2\}}}{2}, \tag{54}$$

or equivalently,

$$\left\langle \phi_{\{1\}} \right\rangle^* - \frac{\overline{\phi}_{\{1\}} + \underline{\phi}_{\{1\}}}{2} = \left\langle \phi_{\{2\}} \right\rangle^* - \frac{\overline{\phi}_{\{2\}} + \underline{\phi}_{\{2\}}}{2}. \quad \Box \tag{55}$$

We remark that for the case $\phi_{\{1\}} = \text{const.}$ and $\phi_{\{2\}} = \text{const.}$, the payoffs and the potential are

$$s_1(\boldsymbol{a}) = \phi_{1,0} + \phi_{\{1,2\}}(a_1, a_2)$$
$$s_2(\boldsymbol{a}) = \phi_{2,0} + \phi_{\{1,2\}}(a_1, a_2) \tag{56}$$

and

$$\Phi(\boldsymbol{a}) = \phi_0 + \phi_{\{1,2\}}(a_1, a_2). \tag{57}$$

Such situation is called as a common interest game. For common interest games, vectors $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{1}\}$ are not linearly independent, where $\boldsymbol{1}$ is the vector of all ones. Additional analysis is needed for investigating the existence of ZD strategies controlling payoffs for common interest games. For example, a two-player two-action common interest game $\boldsymbol{s}_1 = \boldsymbol{s}_2 = \boldsymbol{\Phi} = (1, -1, -1, 1)^\top$ does not contain any ZD strategies controlling payoffs.

We also note that the linear relation (54) may be regarded as an extension of fairness condition $\langle s_1 \rangle^* = \langle s_2 \rangle^*$ in two-player symmetric potential games [42]. Therefore, such ZD strategies may be regarded as an extension of unbeatable strategies in two-player symmetric games to asymmetric cases.

4.2. Asymmetric prisoner's dilemma

As an example, we consider an asymmetric prisoner's dilemma game. The game is described as $N = 2$, $A_1 = A_2 = \{C, D\}$, and

$$\boldsymbol{s}_1 := \left(s_1(C,C), s_1(C,D), s_1(D,C), s_1(D,D) \right)^\top = \left(R_1, S_1, T_1, P_1 \right)^\top$$
$$\boldsymbol{s}_2 := \left(s_2(C,C), s_2(C,D), s_2(D,C), s_2(D,D) \right)^\top = \left(R_2, T_2, S_2, P_2 \right)^\top \tag{58}$$

with $T_j > R_j > P_j > S_j$ $(j = 1, 2)$. (The actions C and D imply cooperation and defection, respectively.) If this game is a potential game, according to Definition 2, a potential function Φ satisfies

$$R_1 - T_1 = \Phi(C,C) - \Phi(D,C)$$
$$S_1 - P_1 = \Phi(C,D) - \Phi(D,D)$$
$$R_2 - T_2 = \Phi(C,C) - \Phi(C,D)$$
$$S_2 - P_2 = \Phi(D,C) - \Phi(D,D). \tag{59}$$

Therefore, for such Φ to exist, the payoffs must satisfy

$$R_1 - T_1 - S_1 + P_1 = R_2 - T_2 - S_2 + P_2. \tag{60}$$

Under this condition, the potential is determined as

$$\Phi(C,C) = \Phi_0$$
$$\Phi(C,D) = \Phi_0 + T_2 - R_2$$
$$\Phi(D,C) = \Phi_0 + T_1 - R_1$$
$$\Phi(D,D) = \Phi_0 + T_2 - R_2 + P_1 - S_1, \tag{61}$$

where Φ_0 is an arbitrary constant.

According to Eqs. (37) and (38),

$$\Phi(\boldsymbol{a}) - s_1(\boldsymbol{a}) = \phi_\emptyset + \phi_{\{2\}}(a_2)$$
$$\Phi(\boldsymbol{a}) - s_2(\boldsymbol{a}) = \phi_\emptyset + \phi_{\{1\}}(a_1). \tag{62}$$

It should be noted that such property that the right-hand sides depend only on the opponent's action is called *coordination-dummy separability* [55]. Then, we obtain

$$\phi_{\{1\}}(C) = \Phi_0 - R_2 - \phi_\emptyset$$
$$\phi_{\{1\}}(D) = \Phi_0 + T_1 - R_1 - S_2 - \phi_\emptyset$$
$$\phi_{\{2\}}(C) = \Phi_0 - R_1 - \phi_\emptyset$$

$$\phi_{\{2\}}(D) = \Phi_0 + T_2 - R_2 - S_1 - \phi_\emptyset. \tag{63}$$

By using the definition (40) and (41), we find $\overline{a}_1 = \overline{a}_2 = D$, $\underline{a}_1 = \underline{a}_2 = C$, and

$$\overline{\phi}_{\{1\}} = \Phi_0 + T_1 - R_1 - S_2 - \phi_\emptyset$$

$$\underline{\phi}_{\{1\}} = \Phi_0 - R_2 - \phi_\emptyset$$

$$\overline{\phi}_{\{2\}} = \Phi_0 + T_2 - R_2 - S_1 - \phi_\emptyset$$

$$\underline{\phi}_{\{2\}} = \Phi_0 - R_1 - \phi_\emptyset. \tag{64}$$

Therefore, according to Theorem 2, a ZD strategy unilaterally enforcing

$$\langle s_1 \rangle^* - \frac{T_1 + S_1}{2} = \langle s_2 \rangle^* - \frac{T_2 + S_2}{2} \tag{65}$$

exists.

We remark that, as shown in the proof of Theorem 2, which player can use the ZD strategy depends on the values $\overline{\phi}_{\{1\}} - \underline{\phi}_{\{1\}}$ and $\overline{\phi}_{\{2\}} - \underline{\phi}_{\{2\}}$. For the ZD strategy of player 1 to exist, the inequality $\overline{\phi}_{\{1\}} - \underline{\phi}_{\{1\}} \geq \overline{\phi}_{\{2\}} - \underline{\phi}_{\{2\}}$ must hold, which is equivalent to

$$R_2 - P_2 \geq R_1 - P_1. \tag{66}$$

Similarly, for the ZD strategy of player 2 to exist, the inequality

$$R_1 - P_1 \geq R_2 - P_2 \tag{67}$$

must hold. The concrete transition probability of the ZD strategy is constructed by using Theorem 1. Explicitly, when the inequality (66) holds, the ZD strategy of player 1 is given by

$$T_1(C) := \begin{pmatrix} T_1(C|C,C) \\ T_1(C|C,D) \\ T_1(C|D,C) \\ T_1(C|D,D) \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \frac{1}{W} \begin{pmatrix} \frac{R_1 - P_1 - R_2 + P_2}{2} \\ \frac{S_1 - T_1 - T_2 + S_2}{2} \\ \frac{T_1 - S_1 - S_2 + T_2}{2} \\ \frac{P_1 - R_1 - P_2 + R_2}{2} \end{pmatrix}, \tag{68}$$

where

$$W = \max \left\{ \left| \frac{R_1 - P_1 - R_2 + P_2}{2} \right|, \left| \frac{S_1 - T_1 - T_2 + S_2}{2} \right| \right\}. \tag{69}$$

It should be noted that, for the symmetric case $(R_1, S_1, T_1, P_1) = (R_2, S_2, T_2, P_2)$, this ZD strategy is reduced to the Tit-for-Tat strategy [1]. The ZD strategy of player 2 when the inequality (67) holds is obtained similarly.

5. Discussion

In this section, we discuss a relation between our Theorem 1 and previous studies. In Ref. [17], McAvoy and Hauert proved the following proposition, which holds even if the action spaces are arbitrary (not limited to finite sets).

Proposition 3 ([17]). *Suppose that there exist two actions* $\underline{a}_j, \overline{a}_j \in A_j$ *and* $W > 0$ *such that*

$$-W \leq B\left(\underline{a}_j, a_{-j}\right) \leq 0 \quad (\forall a_{-j} \in A_{-j})$$

$$0 \leq B\left(\overline{a}_j, a_{-j}\right) \leq W \quad (\forall a_{-j} \in A_{-j}). \tag{70}$$

Then, when we restrict the action set of player j from A_j *to* $A'_j := \left\{ \underline{a}_j, \overline{a}_j \right\}$, *the memory-one strategy of player j*

$$T_j\left(a_j | a'\right) = p\left(a'\right) \delta_{a_j, \underline{a}_j} + \left(1 - p\left(a'\right)\right) \delta_{a_j, \overline{a}_j} \tag{71}$$

with

$$p\left(\underline{a}_j, a'_{-j}\right) := \frac{1}{W} B\left(\underline{a}_j, a_{-j}\right) + 1 \quad (\forall a_{-j} \in A_{-j})$$

$$p\left(\overline{a}_j, a'_{-j}\right) := \frac{1}{W} B\left(\overline{a}_j, a_{-j}\right) \quad (\forall a_{-j} \in A_{-j}) \tag{72}$$

is a ZD strategy unilaterally enforcing $\langle B \rangle^* = 0$.

Indeed, we can find

$$\hat{T}_j \left(\underline{a}_j | \boldsymbol{a}' \right) = \frac{1}{W} B \left(\boldsymbol{a}' \right) \quad \left(\forall a'_j \in A'_j, \forall a'_{-j} \in A_{-j} \right). \tag{73}$$

They called such ZD strategies *two-point autocratic strategies*.

Interestingly, the condition (70) is exactly the same as Eq. (10), which is a necessary and sufficient condition for the existence of ZD strategies when an action space of each player is a finite set [26]. Because two-point autocratic strategies are ZD strategies with A_j consisting of two actions, they satisfy the necessary condition in Ref. [26]. Therefore, the condition (70) is also a necessary and sufficient condition for the existence of two-point autocratic strategies when all action spaces are finite sets. These facts are summarized to the following theorem.

Theorem 3. *If A_k is a finite set for all $k \in \mathcal{N}$, the following three conditions are equivalent.*

 (i) *A ZD strategy of player j controlling B exists.*
 (ii) *A two-point autocratic strategy of player j controlling B exists.*
(iii) *There exist two different actions $\overline{a}_j, \underline{a}_j \in A_j$ of player j such that*

$$B \left(\overline{a}_j, a_{-j} \right) \geq 0 \quad \left(\forall a_{-j} \right)$$
$$B \left(\underline{a}_j, a_{-j} \right) \leq 0 \quad \left(\forall a_{-j} \right), \tag{74}$$

and B is not identically zero.

By definition, two-point autocratic strategies are one-dimensional. However, it should be noted that a strategy (71) is not a ZD strategy in the sense of Definition 1, since Eq. (8) holds not for $\forall \boldsymbol{a}' \in \mathcal{A}$, but for $\forall a'_j \in A'_j, \forall a'_{-j} \in A_{-j}$. Theorem 1 claims that, when a ZD strategy exists, it can be expressed by one-dimensional transition probability using the whole action space A_j. In Ref. [17], it was pointed out that, if we use two-point autocratic strategies, feasible payoff regions are restricted because a two-point autocratic strategy uses only two actions. Our one-dimensional ZD strategies can be useful for bypassing the problem of shrinkage of feasible payoff regions.

6. Concluding remarks

In this paper, we proved two theorems (Theorems 1 and 2) which can be useful for implementation of ZD strategies in repeated games. The first theorem shows that ZD strategies, if exist, can be realized by one-dimensional transition probability. This result can be useful when an action space has one-dimensional structure. The second theorem claims that ZD strategies controlling payoffs exist in two-player non-trivial potential games. We explained this result in the asymmetric prisoner's dilemma game as an example.

Before ending this paper, we make two remarks. The first remark is related to extension of Theorem 2 to the case $N > 2$. Since the proof of Theorem 2 relies on the fact $N = 2$, we cannot straightforwardly extend its proof to $N > 2$ case. Existence of ZD strategies controlling payoffs in potential games may be a specific result in $N = 2$ case. Further investigation is needed to clarify whether extension of our result to $N > 2$ is possible or not.

The second remark is related to the implication of Theorem 1. The form of the transition probability (25) is simple, and, as discussed in Section 3, each element depends only on the quantity at the previous action profile $\boldsymbol{a}'$. Although this property is similar to that of the trap model, the stationary probability distribution cannot be written as a simple form, because the total dynamics is determined by strategies of all players. We would like to investigate mathematical properties of the transition probability (25) in future. Furthermore, specifying situations where one-dimensional implementation (25) of a ZD strategy is necessary is an interesting future problem.

Data availability

No data was used for the research described in the article.

References

[1] W.H. Press, F.J. Dyson, Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent, Proc. Natl. Acad. Sci. 109 (26) (2012) 10409–10413.
[2] A.J. Stewart, J.B. Plotkin, From extortion to generosity, evolution in the iterated prisoner's dilemma, Proc. Natl. Acad. Sci. 110 (38) (2013) 15348–15353.

[3] C. Hilbe, M.A. Nowak, K. Sigmund, Evolution of extortion in iterated prisoner's dilemma games, Proc. Natl. Acad. Sci. 110 (17) (2013) 6913–6918.

[4] C. Adami, A. Hintze, Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything, Nat. Commun. 4 (1) (2013) 2193.

[5] C. Hilbe, M.A. Nowak, A. Traulsen, Adaptive dynamics of extortion and compliance, PLoS ONE 8 (11) (2013) e77886.

[6] E. Akin, The iterated prisoner's dilemma: good strategies and their dynamics, in: Ergodic Theory in Advances in Dynamical Systems, 2016, pp. 77–107.

[7] X. Chen, L. Wang, F. Fu, The intricate geometry of zero-determinant strategies underlying evolutionary adaptation from extortion to generosity, New J. Phys. 24 (10) (2022) 103001.

[8] X. Chen, F. Fu, Outlearning extortioners: unbending strategies can Foster reciprocal fairness and cooperation, PNAS Nexus 2 (6) (2023) pgad176.

[9] A. Szolnoki, M. Perc, Evolution of extortion in structured populations, Phys. Rev. E 89 (2) (2014) 022804.

[10] A. Szolnoki, M. Perc, Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations, Sci. Rep. 4 (1) (2014) 5496.

[11] Z. Chen, Y. Geng, X. Chen, F. Fu, Unbending strategies shepherd cooperation and suppress extortion in spatial populations, New J. Phys. 26 (7) (2024) 073047.

[12] C. Hilbe, T. Röhl, M. Milinski, Extortion subdues human players but is finally punished in the prisoner's dilemma, Nat. Commun. 5 (2014) 3976.

[13] Z. Wang, Y. Zhou, J.W. Lien, J. Zheng, B. Xu, Extortion can outperform generosity in the iterated prisoner's dilemma, Nat. Commun. 7 (2016) 11125.

[14] L. Becks, M. Milinski, Extortion strategies resist disciplining when higher competitiveness is rewarded with extra gain, Nat. Commun. 10 (1) (2019) 783.

[15] C. Hilbe, B. Wu, A. Traulsen, M.A. Nowak, Cooperation and control in multiplayer social dilemmas, Proc. Natl. Acad. Sci. 111 (46) (2014) 16425–16430.

[16] L. Pan, D. Hao, Z. Rong, T. Zhou, Zero-determinant strategies in iterated public goods game, Sci. Rep. 5 (2015) 13096.

[17] A. McAvoy, C. Hauert, Autocratic strategies for iterated games with arbitrary action spaces, Proc. Natl. Acad. Sci. 113 (13) (2016) 3573–3578.

[18] C. Hilbe, A. Traulsen, K. Sigmund, Partners or rivals? Strategies for the iterated prisoner's dilemma, Games Econ. Behav. 92 (2015) 41–52.

[19] G. Ichinose, N. Masuda, Zero-determinant strategies in finitely repeated games, J. Theor. Biol. 438 (2018) 61–77.

[20] A. Govaert, M. Cao, Zero-determinant strategies in repeated multiplayer social dilemmas with discounted payoffs, IEEE Trans. Autom. Control 66 (10) (2020) 4575–4588.

[21] D. Hao, Z. Rong, T. Zhou, Extortion under uncertainty: zero-determinant strategies in noisy games, Phys. Rev. E 91 (2015) 052803.

[22] M. Ueda, T. Tanaka, Linear algebraic structure of zero-determinant strategies in repeated games, PLoS ONE 15 (4) (2020) e0230973.

[23] A. Mamiya, G. Ichinose, Zero-determinant strategies under observation errors in repeated games, Phys. Rev. E 102 (2020) 032115.

[24] A. McAvoy, C. Hauert, Autocratic strategies for alternating games, Theor. Popul. Biol. 113 (2017) 13–22.

[25] X. He, H. Dai, P. Ning, R. Dutta, Zero-determinant strategies for multi-player multi-action iterated games, IEEE Signal Process. Lett. 23 (3) (2016) 311–315.

[26] M. Ueda, Necessary and sufficient condition for the existence of zero-determinant strategies in repeated games, J. Phys. Soc. Jpn. 91 (8) (2022) 084801.

[27] M. Ueda, Tit-for-tat strategy as a deformed zero-determinant strategy in repeated games, J. Phys. Soc. Jpn. 90 (2) (2021) 025002.

[28] M. Ueda, Memory-two zero-determinant strategies in repeated games, R. Soc. Open Sci. 8 (5) (2021) 202186.

[29] M. Ueda, Controlling conditional expectations by zero-determinant strategies, Oper. Res. Forum 3 (3) (2022) 48.

[30] A. Al Daoud, G. Kesidis, J. Liebeherr, Zero-determinant strategies: a game-theoretic approach for sharing licensed spectrum bands, IEEE J. Sel. Areas Commun. 32 (11) (2014) 2297–2308.

[31] H. Zhang, D. Niyato, L. Song, T. Jiang, Z. Han, Zero-determinant strategy for resource sharing in wireless cooperations, IEEE Trans. Wirel. Commun. 15 (3) (2016) 2179–2192.

[32] A. Farraj, E. Hammad, A. Al Daoud, D. Kundur, A game-theoretic analysis of cyber switching attacks and mitigation in smart grid systems, IEEE Trans. Smart Grid 7 (4) (2015) 1846–1855.

[33] C. Tang, C. Li, X. Yu, Z. Zheng, Z. Chen, Cooperative mining in blockchain networks with zero-determinant strategies, IEEE Trans. Cybern. 50 (10) (2019) 4544–4549.

[34] H. Shi, S. Wang, Q. Hu, X. Cheng, J. Zhang, J. Yu, Fee-free pooled mining for countering pool-hopping attack in blockchain, IEEE Trans. Dependable Secure Comput. 18 (4) (2020) 1580–1590.

[35] C. Tang, X. Li, M. Cao, Z. Zhang, X. Yu, Incentive mechanism for macrotasking crowdsourcing: a zero-determinant strategy approach, IEEE Int. Things J. 6 (5) (2019) 8589–8601.

[36] Q. Hu, S. Wang, P. Ma, X. Cheng, W. Lv, R. Bie, Quality control in crowdsourcing using sequential zero-determinant strategies, IEEE Trans. Knowl. Data Eng. 32 (5) (2019) 998–1009.

[37] Q. Hu, S. Wang, X. Cheng, J. Zhang, W. Lv, Cost-efficient mobile crowdsensing with spatial-temporal awareness, IEEE Trans. Mob. Comput. 20 (3) (2019) 928–938.

[38] S. Wang, H. Shi, Q. Hu, B. Lin, X. Cheng, Moving target defense for Internet of things based on the zero-determinant theory, IEEE Int. Things J. 7 (1) (2019) 661–668.

[39] D. Zhang, Y. Tian, C. Yue, M. Fan, Cooperate delegation of computation for rational party using zero-determinant strategy approach, IEEE Access 8 (2020) 27734–27743.

[40] Q. Hu, S. Wang, Z. Xiong, X. Cheng, Nothing wasted: full contribution enforcement in federated edge learning, IEEE Trans. Mob. Comput. 22 (2021) 2850–2861.

[41] S. Wang, L. Shi, Q. Hu, J. Zhang, X. Cheng, J. Yu, Privacy-aware data trading, IEEE Trans. Inf. Forensics Secur. 16 (2021) 3916–3927.

[42] M. Ueda, Unbeatable tit-for-tat as a zero-determinant strategy, J. Phys. Soc. Jpn. 91 (5) (2022) 054804.

[43] M. Ueda, Unexploitable games and unbeatable strategies, IEEE Access 11 (2023) 5062–5068.

[44] A. Rapoport, A.M. Chammah, C.J. Orwant, Prisoner's Dilemma: A Study in Conflict and Cooperation, vol. 165, University of Michigan Press, 1965.

[45] R. Axelrod, W.D. Hamilton, The evolution of cooperation, Science 211 (4489) (1981) 1390–1396.

[46] P. Duersch, J. Oechssler, B.C. Schipper, When is tit-for-tat unbeatable?, Int. J. Game Theory 43 (1) (2014) 25–36.

[47] P. Duersch, J. Oechssler, B.C. Schipper, Unbeatable imitation, Games Econ. Behav. 76 (1) (2012) 88–96.

[48] C. Xia, J. Wang, M. Perc, Z. Wang, Reputation and reciprocity, Phys. Life Rev. 46 (2023) 8–45.

[49] V. Capraro, R. Di Paolo, M. Perc, V. Pizziol, Language-based game theory in the age of artificial intelligence, J. R. Soc. Interface 21 (212) (2024) 20230720.

[50] D. Fudenberg, J. Tirole, Game Theory, MIT Press, Massachusetts, 1991.

[51] M.J. Osborne, A. Rubinstein, A Course in Game Theory, MIT Press, Massachusetts, 1994.

[52] D. Monderer, L.S. Shapley, Potential games, Games Econ. Behav. 14 (1) (1996) 124–143.

[53] T. Ui, A Shapley value representation of potential games, Games Econ. Behav. 31 (1) (2000) 121–135.

[54] J.-P. Bouchaud, A. Georges, Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications, Phys. Rep. 195 (4–5) (1990) 127–293.

[55] Q.D. Lã, Y.H. Chew, B.-H. Soong, Potential Game Theory, Springer, 2016.