

Unbeatable Tit-for-Tat as a Zero-Determinant Strategy

Masahiko Ueda^{1*}

¹*Graduate School of Sciences and Technology for Innovation, Yamaguchi University,
Yamaguchi 753-8511, Japan*

Tit-for-Tat strategy is a strategy in repeated two-player symmetric games which imitates the previous action of the opponent. We show that the Tit-for-Tat strategy is a zero-determinant strategy, which unilaterally equalizes the expected payoffs of two players, if and only if the stage game is a potential game. Because it has been known that this condition is equivalent to the condition that the Tit-for-Tat strategy is unbeatable, our results suggest some relation between unbeatable property and the concept of zero-determinant strategy.

1. Introduction

The Tit-for-Tat (TFT) strategy was discovered as a cooperative strategy in the infinitely repeated prisoner's dilemma game, which imitates the previous action of the opponent.¹⁾ Although a pair of TFT does not form subgame perfect equilibrium, it forms Nash equilibrium. Axelrod obtained the numerical results that TFT is the most successful strategy in the prisoner's dilemma game by using computer tournaments.²⁾ Meanwhile, in evolutionary games, it was pointed out that TFT is not successful because it is not robust against errors.³⁻⁵⁾ Recently, it was found that TFT is contained in the class of zero-determinant (ZD) strategies, which unilaterally enforce linear relations between expected payoffs.⁶⁾ Furthermore, it was shown that TFT is also a deformed ZD strategy which unilaterally equalizes all moments of payoffs of two players.⁷⁾ Variants of TFT were recently proposed which are robust against implementation errors.^{8,9)}

Although many researchers investigated properties of TFT in the repeated prisoner's dilemma games, little is known about properties of TFT in other repeated two-player symmetric games. Although simple, imitation strategies are generally successful in several situations.¹⁰⁻¹³⁾ Recently, Duersch et al. found that TFT is unbeatable if and only if the stage game is a potential game.¹⁴⁾ Potential games are a class of games in strategic form which

*m.ueda@yamaguchi-u.ac.jp

have potential functions.¹⁵⁾ For potential games, Nash equilibrium is realized as the maximum of a potential. Potential games contain several important situations such as the Cournot duopoly game and the public goods game, in addition to the prisoner's dilemma game. For non-potential games, such as rock-paper-scissors game, TFT can be exploited unboundedly.

In this paper, we show that, in infinitely repeated two-player symmetric games, TFT is unbeatable if and only if TFT is a ZD strategy. ZD strategies attract much attention because of their counterintuitive properties that the payoffs are unilaterally controlled by one player.^{16–20)} We show that TFT is a ZD strategy, which unilaterally equalizes the expected payoffs of two players, if and only if the stage game is a potential games, even if the stage game is not the prisoner's dilemma game. When combined with the results of Ref.,¹⁴⁾ we can see that the unbeatable property of TFT is equivalent to that TFT is a ZD strategy.

This paper is organized as follows. In Section 2, we introduce a model of infinitely repeated two-player symmetric games. In Section 3, we introduce basic concepts used in the later sections and results of the previous papers.^{6,15)} In Section 4, we prove our main theorem that TFT is a ZD strategy if and only if the stage game is a potential game. In addition, we also show that TFT unilaterally equalizes the expected payoffs of two players in potential games. Moreover, we show that TFT cannot unilaterally enforce any linear relations between expected payoffs in non-potential games, if the opponent uses memory-one strategies. In Section 5, we check the main result in two examples. In Section 6, we introduce the results of Ref.,¹⁴⁾ and discuss the relation between our results and the results of Ref.¹⁴⁾ In this section, we also provide the results about other imitation strategies. Section 7 is devoted to concluding remarks.

2. Model

We consider a two-player symmetric game. The set of player is $N := \{1, 2\}$. The set of action of player a in the stage game is $A_a = A := \{1, \dots, M\}$, where M is a natural number representing the number of action. The action of player a is written as $\sigma_a \in A$. We collectively write $\sigma := (\sigma_1, \sigma_2)$, and call σ a state. The payoff of player $a \in \{1, 2\}$ in the stage game when the state is σ is described as $s_a(\sigma)$. Therefore, the stage game is described as $G := (N, \{A_a\}_{a \in N}, \{s_a\}_{a \in N})$.²¹⁾ We introduce the notation that $-a := N \setminus \{a\}$. We assume that the game is symmetric, that is, $s_2(\sigma_1, \sigma_2) = s_1(\sigma_2, \sigma_1) \quad (\forall \sigma_1, \forall \sigma_2)$.

We repeat the stage game G infinitely. We write an action of player a at round $t \geq 1$ as $\sigma_a^{(t)}$. We also introduce the notation $h_{[t:t']}] := (\sigma^{(t)}, \dots, \sigma^{(t')})$ for $t \leq t'$, and call $h_{[t:t']}$ the history in time interval $[t : t']$. A strategy of player a in the infinitely repeated game is defined

by $\left\{T_a^{(t)}\left(\sigma_a^{(t)}|h_{[1:t-1]}\right)\right\}_{t=1}^{\infty}$, where $T_a^{(t)}\left(\sigma_a^{(t)}|h_{[1:t-1]}\right)$ is the conditional probability of taking action $\sigma_a^{(t)}$ at round t when the history is $h_{[1:t-1]}$. We write the expectation of the quantity B with respect to strategies of both players by $\mathbb{E}[B]$. The payoff of player a in the infinitely repeated game is defined by

$$\mathcal{S}_a := (1 - \delta)\mathbb{E}\left[\sum_{t=1}^{\infty}\delta^{t-1}s_a\left(\boldsymbol{\sigma}^{(t)}\right)\right], \quad (1)$$

where δ is a discounting factor satisfying $0 \leq \delta \leq 1$.

Below we consider only the case $\delta = 1$, where the payoff of player a is described as

$$\mathcal{S}_a = \lim_{T \rightarrow \infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^T s_a\left(\boldsymbol{\sigma}^{(t)}\right)\right]. \quad (2)$$

3. Preliminaries

In this section, we introduce several concepts used in later sections. Below, the quantity $\delta_{\boldsymbol{\sigma}, \boldsymbol{\sigma}'}$ represents the Kronecker delta. We also define $s_0(\boldsymbol{\sigma}) := 1 (\forall \boldsymbol{\sigma})$.

First, we introduce time-independent memory- n strategies.

Definition 1 *A strategy of player a is a time-independent memory- n strategy ($n \geq 0$) when it is written in the form*

$$T_a^{(t)}\left(\sigma_a^{(t)}|h_{[1:t-1]}\right) = T_a\left(\sigma_a^{(t)}|h_{[t-n:t-1]}\right) \quad (\forall \sigma_a^{(t)}, \forall h_{[1:t-1]}) \quad (3)$$

for all $t \geq n + 1$ with some common conditional probability T_a .

It should be noted that, in order to define a strategy, the initial condition for $t \leq n$ must also be given aside from T_a .

As a special time-independent memory-one strategy, we introduce the Tit-for-Tat strategy.

Definition 2 *A time-independent memory-one strategy of player a is the Tit-for-Tat (TFT) strategy when T_a in Definition 1 is written in the form*

$$T_a(\sigma_a|\boldsymbol{\sigma}') = \delta_{\sigma_a, \sigma'_a} \quad (\forall \sigma_a, \forall \boldsymbol{\sigma}'). \quad (4)$$

That is, TFT imitates the action of the opponent in the previous round.

Next, we introduce zero-determinant strategies. For time-independent memory-one strategies T_a of player a , we first introduce the Press-Dyson vectors^{22,23)}

$$\hat{T}_a(\sigma_a|\boldsymbol{\sigma}') := T_a(\sigma_a|\boldsymbol{\sigma}') - \delta_{\sigma_a, \sigma'_a} \quad (\forall \sigma_a, \forall \boldsymbol{\sigma}'). \quad (5)$$

Because the second term in the right-hand side of Eq. (5) can be regarded as the strategy ‘‘Repeat’’, which repeats his/her own action in the previous round, the Press-Dyson vectors are interpreted as the difference between his/her own strategy and ‘‘Repeat’’. By using the

Press-Dyson vectors, we define the zero-determinant strategies.

Definition 3 *A time-independent memory-one strategy of player a is a zero-determinant (ZD) strategy when its Press-Dyson vectors can be written in the form*

$$\sum_{\sigma_a} c_{\sigma_a} \hat{T}_a(\sigma_a | \sigma') = \sum_{b=0}^2 \alpha_b s_b(\sigma') \quad (\forall \sigma') \quad (6)$$

with some nontrivial coefficients $\{\alpha_b\}$ and $\{c_{\sigma_a}\}$ (that is, not $\alpha_0 = \alpha_1 = \dots = \alpha_N = 0$, and not $c_1 = \dots = c_M = \text{const.}$).

In other words, in ZD strategies, a linear combination of the Press-Dyson vectors is described as a linear combination of payoff vectors and a vector of all ones. We remark that the definition of ZD strategies of player a does not depend on the length of memory of strategies of player $-a$.

In order to see properties of ZD strategies, we first remember that the joint probability of states satisfies the recursion relation

$$P(h_{[1:t+1]}) = \left\{ \prod_a T_a^{(t+1)}(\sigma_a^{(t+1)} | h_{[1:t]}) \right\} P(h_{[1:t]}). \quad (7)$$

We also define probability distribution of $\sigma^{(t)}$ by

$$P_t(\sigma^{(t)}) := \sum_{h_{[1:t-1]}} P(h_{[1:t]}). \quad (8)$$

We consider the situation that player a uses a ZD strategy. By taking $\sum_{\sigma_{-a}^{(t+1)}} \sum_{h_{[1:t]}}$ in both sides of Eq. (7), we obtain

$$\sum_{\sigma'} \delta_{\sigma_a, \sigma_a^{(t+1)}} P_{t+1}(\sigma') = \sum_{\sigma'} T_a(\sigma_a^{(t+1)} | \sigma') P_t(\sigma') \quad (9)$$

Then, by replacing $\sigma_a^{(t+1)} \rightarrow \sigma_a$ and calculating $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T$ of both sides, we obtain

$$\sum_{\sigma'} \delta_{\sigma_a, \sigma_a} P^*(\sigma') = \sum_{\sigma'} T_a(\sigma_a | \sigma') P^*(\sigma'), \quad (10)$$

where we have defined the limit distribution

$$P^*(\sigma) := \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T P_t(\sigma). \quad (11)$$

This fact is known as Akin's lemma:

Lemma 1 ^(22,23) *The Press-Dyson vectors satisfy*

$$\sum_{\sigma'} P^*(\sigma') \hat{T}_a(\sigma_a | \sigma') = 0 \quad (\forall \sigma_a). \quad (12)$$

We also remark that the payoffs in the repeated games are described by expectation with

respect to the limit distribution:

$$\begin{aligned} \mathcal{S}_b &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \sum_{\sigma} s_b(\sigma) P_t(\sigma) \\ &= \sum_{\sigma} s_b(\sigma) P^*(\sigma). \end{aligned} \quad (13)$$

Below we write the expectation of quantity B with respect to the limit distribution P^* as $\langle B \rangle^*$. Therefore, $\mathcal{S}_b = \langle s_b \rangle^*$. The following proposition about ZD strategies is a direct consequence of Akin's lemma.

Proposition 1 ^(6,23) *A ZD strategy (6) unilaterally enforces a linear relation between expected payoffs:*

$$0 = \sum_{b=0}^2 \alpha_b \langle s_b \rangle^*. \quad (14)$$

In other words, ZD strategies unilaterally control expected payoffs. We emphasize that this property of ZD strategies hold regardless of the strategy of player $-a$. Although we above consider the situation that the action set A is countable, ZD strategies were also extended to games with uncountable action set.²⁴⁾ For such cases, the argument about probability is replaced by that about probability density, and the Kronecker delta is replaced by the Dirac delta. Furthermore, we also remark that the concept of ZD strategies was recently extended to memory- n strategies with $n \geq 1$.^{20,25)}

Finally, we introduce the concept of potential game.¹⁵⁾

Definition 4 *A game $G = (N, \{A_a\}_{a \in N}, \{s_a\}_{a \in N})$ is an (exact) potential game when there exist a common function $\Phi(\sigma)$ satisfying*

$$s_a(\sigma_a, \sigma_{-a}) - s_a(\sigma'_a, \sigma_{-a}) = \Phi(\sigma_a, \sigma_{-a}) - \Phi(\sigma'_a, \sigma_{-a}) \quad (\forall \sigma_a, \forall \sigma'_a, \forall \sigma_{-a}) \quad (15)$$

for all player a .

The function Φ is called a potential function. Because the Nash equilibrium σ^* is defined by the condition

$$s_a(\sigma_a^*, \sigma_{-a}^*) \geq s_a(\sigma_a, \sigma_{-a}^*) \quad (\forall a, \forall \sigma_a), \quad (16)$$

the condition of the Nash equilibrium for potential games is rewritten as

$$\Phi(\sigma_a^*, \sigma_{-a}^*) \geq \Phi(\sigma_a, \sigma_{-a}^*) \quad (\forall a, \forall \sigma_a). \quad (17)$$

Therefore, for a potential game, the Nash equilibrium is realized as the maximum of a potential function. It should be remarked that the concept of potential game is also defined for the case that the action set A is uncountable.

4. Results

In the prisoner's dilemma game, it is known that TFT is a ZD strategy, which unilaterally enforces $\langle s_1 \rangle^* = \langle s_2 \rangle^*$.⁶⁾ A natural question is “Is TFT also a ZD strategy in other two-player symmetric games?”. In this section, we show that TFT in two-player symmetric games becomes a ZD strategy if and only if the stage game is a potential game.

We consider the situation that player 1 takes TFT. Below, for quantities $B(\sigma_1, \sigma_2)$, we use the following notations

$$B^{(S)}(\sigma_1, \sigma_2) := \frac{1}{2} [B(\sigma_1, \sigma_2) + B(\sigma_2, \sigma_1)] \quad (18)$$

$$B^{(A)}(\sigma_1, \sigma_2) := \frac{1}{2} [B(\sigma_1, \sigma_2) - B(\sigma_2, \sigma_1)], \quad (19)$$

which correspond to symmetric and anti-symmetric parts of $B(\sigma_1, \sigma_2)$, respectively.

4.1 When is TFT a zero-determinant strategy?

We first prove the following lemma, which is essentially the same as one in Ref.²⁶⁾

Lemma 2 *For two-player symmetric games, the definition of potential game is equivalent to the condition*

$$s_1^{(A)}(\sigma_1, \sigma_2) = c_{\sigma_2} - c_{\sigma_1} \quad (\forall \sigma_1, \forall \sigma_2) \quad (20)$$

with some function c_σ .

Proof. For two-player symmetric game, the definition of potential game is explicitly written as

$$s_1(\sigma_1, \sigma_2) - s_1(\sigma'_1, \sigma_2) = \Phi(\sigma_1, \sigma_2) - \Phi(\sigma'_1, \sigma_2) \quad (\forall \sigma_1, \forall \sigma'_1, \forall \sigma_2) \quad (21)$$

$$s_2(\sigma_1, \sigma_2) - s_2(\sigma_1, \sigma'_2) = \Phi(\sigma_1, \sigma_2) - \Phi(\sigma_1, \sigma'_2) \quad (\forall \sigma_2, \forall \sigma'_2, \forall \sigma_1). \quad (22)$$

Because the game is symmetric, the condition (22) is equivalent to

$$s_1(\sigma_2, \sigma_1) - s_1(\sigma'_2, \sigma_1) = \Phi(\sigma_1, \sigma_2) - \Phi(\sigma_1, \sigma'_2). \quad (23)$$

By relabeling the name of variables, it is rewritten as

$$s_1(\sigma_1, \sigma_2) - s_1(\sigma'_1, \sigma_2) = \Phi(\sigma_2, \sigma_1) - \Phi(\sigma_2, \sigma'_1). \quad (24)$$

Then we obtain

$$\Phi(\sigma_1, \sigma_2) - \Phi(\sigma'_1, \sigma_2) = \Phi(\sigma_2, \sigma_1) - \Phi(\sigma_2, \sigma'_1), \quad (25)$$

or

$$\Phi^{(A)}(\sigma_1, \sigma_2) = \Phi^{(A)}(\sigma'_1, \sigma_2), \quad (26)$$

which means that the anti-symmetric part of the potential Φ does not depend on σ_1 . By using the same argument, we also obtain

$$\Phi^{(A)}(\sigma_1, \sigma_2) = \Phi^{(A)}(\sigma_1, \sigma'_2). \quad (27)$$

Therefore, $\Phi^{(A)}$ must be constant. However, because $\Phi^{(A)}$ is the anti-symmetric part, $\Phi^{(A)}(\sigma, \sigma) = 0$ for any σ , and this constant must be zero. Thus, $\Phi^{(A)}(\sigma_1, \sigma_2) = 0$ for all (σ_1, σ_2) , and we conclude that the potential Φ is symmetric.

By using this fact, we find that

$$\begin{aligned} s_1(\sigma_1, \sigma_2) - s_1(\sigma_2, \sigma_1) &= s_1(\sigma_1, \sigma_2) - s_2(\sigma_1, \sigma_2) \\ &= [\Phi(\sigma_1, \sigma_2) - \Phi(1, \sigma_2) + s_1(1, \sigma_2)] - [\Phi(\sigma_1, \sigma_2) - \Phi(\sigma_1, 1) + s_2(\sigma_1, 1)] \\ &= [-\Phi(1, \sigma_2) + s_1(1, \sigma_2)] - [-\Phi(1, \sigma_1) + s_1(1, \sigma_1)] \\ &= d_{\sigma_2} - d_{\sigma_1}, \end{aligned} \quad (28)$$

where we have defined

$$d_\sigma := -\Phi(1, \sigma) + s_1(1, \sigma). \quad (29)$$

Therefore, we obtain the form (20).

Conversely, when the condition (20) holds,

$$s_1(\sigma_1, \sigma_2) - c_{\sigma_2} = s_1(\sigma_2, \sigma_1) - c_{\sigma_1}. \quad (30)$$

When we introduce the quantity

$$\phi(\sigma_1, \sigma_2) := s_1(\sigma_1, \sigma_2) - c_{\sigma_2}, \quad (31)$$

it satisfies

$$\phi(\sigma_1, \sigma_2) = \phi(\sigma_2, \sigma_1). \quad (32)$$

Then we find that

$$\begin{aligned} s_1(\sigma_1, \sigma_2) - s_1(\sigma'_1, \sigma_2) &= [\phi(\sigma_1, \sigma_2) + c_{\sigma_2}] - [\phi(\sigma'_1, \sigma_2) + c_{\sigma_2}] \\ &= \phi(\sigma_1, \sigma_2) - \phi(\sigma'_1, \sigma_2) \end{aligned} \quad (33)$$

and

$$\begin{aligned} s_2(\sigma_1, \sigma_2) - s_2(\sigma_1, \sigma'_2) &= s_1(\sigma_2, \sigma_1) - s_1(\sigma'_2, \sigma_1) \\ &= \phi(\sigma_2, \sigma_1) - \phi(\sigma'_2, \sigma_1) \\ &= \phi(\sigma_1, \sigma_2) - \phi(\sigma_1, \sigma'_2). \end{aligned} \quad (34)$$

Therefore, the quantity ϕ can be regarded as a potential. \square

We now prove our main theorem.

Theorem 1 *For two-player symmetric games, TFT is a ZD strategy if and only if the stage game is a potential game.*

Proof. When player 1 takes TFT, her Press-Dyson vector is described as

$$\hat{T}_1(\sigma_1|\sigma'_1, \sigma'_2) = \delta_{\sigma_1, \sigma'_2} - \delta_{\sigma_1, \sigma'_1}. \quad (35)$$

If TFT is a ZD strategy (6), it satisfies

$$c_{\sigma'_2} - c_{\sigma'_1} = \sum_{b=1}^2 \alpha_b s_b(\sigma'_1, \sigma'_2) + \alpha_0 \quad (36)$$

with some non-trivial coefficients $\{\alpha_b\}$ and $\{c_{\sigma'_a}\}$. Because

$$c_{\sigma'_1} - c_{\sigma'_2} = \sum_{b=1}^2 \alpha_b s_b(\sigma'_2, \sigma'_1) + \alpha_0, \quad (37)$$

we obtain

$$-\sum_{b=1}^2 \alpha_b s_b(\sigma'_1, \sigma'_2) - \alpha_0 = \sum_{b=1}^2 \alpha_b s_b(\sigma'_2, \sigma'_1) + \alpha_0, \quad (38)$$

or

$$0 = \sum_{b=1}^2 \alpha_b s_b^{(S)}(\sigma'_1, \sigma'_2) + \alpha_0, \quad (39)$$

for all (σ'_1, σ'_2) . Furthermore, because the game is symmetric, this equation can be rewritten as

$$0 = (\alpha_1 + \alpha_2) s_1^{(S)}(\sigma'_1, \sigma'_2) + \alpha_0, \quad (40)$$

Then, the coefficients must satisfy either of the following relations:

$$\begin{cases} \alpha_0 = -(\alpha_1 + \alpha_2) s_1^{(S)}(1, 1) & (\text{if } s_1^{(S)}(\sigma'_1, \sigma'_2) = s_1^{(S)}(1, 1) \quad (\forall \sigma'_1, \forall \sigma'_2)) \\ \alpha_1 + \alpha_2 = 0, \quad \alpha_0 = 0 & (\text{otherwise}). \end{cases} \quad (41)$$

For the former case, Eq. (36) can be rewritten as

$$\begin{aligned} c_{\sigma'_2} - c_{\sigma'_1} &= \sum_{b=1}^2 \alpha_b s_b(\sigma'_1, \sigma'_2) - (\alpha_1 + \alpha_2) s_1^{(S)}(1, 1) \\ &= \alpha_1 [s_1^{(S)}(\sigma'_1, \sigma'_2) + s_1^{(A)}(\sigma'_1, \sigma'_2)] + \alpha_2 [s_1^{(S)}(\sigma'_2, \sigma'_1) + s_1^{(A)}(\sigma'_2, \sigma'_1)] - (\alpha_1 + \alpha_2) s_1^{(S)}(1, 1) \\ &= \alpha_1 s_1^{(A)}(\sigma'_1, \sigma'_2) + \alpha_2 s_1^{(A)}(\sigma'_2, \sigma'_1) \\ &= (\alpha_1 - \alpha_2) s_1^{(A)}(\sigma'_1, \sigma'_2). \end{aligned} \quad (42)$$

For the latter case, Eq. (36) can be rewritten as

$$\begin{aligned}
 c_{\sigma'_2} - c_{\sigma'_1} &= \alpha_1 s_1(\sigma'_1, \sigma'_2) - \alpha_1 s_2(\sigma'_1, \sigma'_2) \\
 &= \alpha_1 s_1(\sigma'_1, \sigma'_2) - \alpha_1 s_1(\sigma'_2, \sigma'_1) \\
 &= 2\alpha_1 s_1^{(A)}(\sigma'_1, \sigma'_2).
 \end{aligned} \tag{43}$$

Therefore, $c_{\sigma'_2} - c_{\sigma'_1}$, which is non-zero for some pairs (σ'_1, σ'_2) , is proportional to $s_1^{(A)}(\sigma'_1, \sigma'_2)$ for both cases. From Lemma 2, this is the condition for a game to be a potential game.

Therefore, if TFT is a ZD strategy, then the game must be a potential game.

Conversely, if the game is a potential game, it satisfies

$$c_{\sigma'_2} - c_{\sigma'_1} = s_1^{(A)}(\sigma'_1, \sigma'_2) \tag{44}$$

for some function c_{σ} (Lemma 2). This means that

$$\begin{aligned}
 c_{\sigma'_2} - c_{\sigma'_1} &= \frac{1}{2} s_1(\sigma'_1, \sigma'_2) - \frac{1}{2} s_1(\sigma'_2, \sigma'_1) \\
 &= \frac{1}{2} s_1(\sigma'_1, \sigma'_2) - \frac{1}{2} s_2(\sigma'_1, \sigma'_2)
 \end{aligned} \tag{45}$$

with

$$c_{\sigma'_2} - c_{\sigma'_1} = \sum_{\sigma_1} c_{\sigma_1} \hat{T}_1(\sigma_1 | \sigma'_1, \sigma'_2). \tag{46}$$

Therefore, if the game is a potential game, TFT is a ZD strategy. \square

The following corollary is a direct consequence of Theorem 1 and Proposition 1.

Corollary 1 *For two-player symmetric games, TFT unilaterally enforces*

$$\langle s_1 \rangle^* = \langle s_2 \rangle^* \tag{47}$$

for potential games.

Proof. In the proof of Theorem 1, we find that

$$\sum_{\sigma_1} c_{\sigma_1} \hat{T}_1(\sigma_1 | \sigma'_1, \sigma'_2) = \frac{1}{2} s_1(\sigma'_1, \sigma'_2) - \frac{1}{2} s_2(\sigma'_1, \sigma'_2) \tag{48}$$

for potential games. By using Proposition 1, we obtain Eq. (47). \square

We remark that the linear relation enforced by TFT in potential games is restricted to Eq. (47). Therefore, TFT can only unilaterally equalizes the expected payoffs of two players in potential games.

4.2 TFT in non-potential games

Theorem 1 claims that TFT cannot be a ZD strategy in non-potential games. However, this does not directly lead to the statement that TFT cannot unilaterally enforce any linear relations between expected payoffs in non-potential games, since there is no guarantee that ZD strategies are only strategies which unilaterally enforce linear relations between expected payoffs. In Ref.,²⁷⁾ the authors proved that memory-one strategies which unilaterally enforce linear relations between expected payoffs are restricted to ZD strategies and unconditional strategies in the prisoner's dilemma game, if both players use memory-one strategies and a stationary distribution of the induced Markov chain exists. Here we extend their results to our case.

Theorem 2 *For two-player symmetric games, if the opponent $-a$ also uses a time-independent memory-one strategy $T_{-a}(\sigma_{-a}|\sigma')$ and the induced Markov chain*

$$P_{t+1}(\sigma) = \sum_{\sigma'} \left\{ \prod_{a=1}^2 T_a(\sigma_a|\sigma') \right\} P_t(\sigma') \quad (49)$$

has a stationary distribution, then the following two conditions are equivalent:

(a) *TFT is a ZD strategy.*

(b) *TFT unilaterally enforces a linear relation between expected payoffs.*

Proof. ((a) \Rightarrow (b)): If TFT is a ZD strategy, Theorem 1 claims that the stage game is a potential game. Then, under the assumptions, Corollary 1 claims that TFT unilaterally enforces a linear relation (47).

((b) \Rightarrow (a)): We first note that the limit distribution P^* coincides with the stationary distribution of Eq. (49) under the assumptions. We assume that TFT of player 1 unilaterally enforces a linear relation between expected payoffs

$$0 = \alpha_1 \langle s_1 \rangle^* + \alpha_2 \langle s_2 \rangle^* + \alpha_0. \quad (50)$$

Due to the assumptions, the stationary distribution exists, such that

$$P^*(\sigma) = \sum_{\sigma'} T(\sigma|\sigma') P^*(\sigma') \quad (51)$$

with the transition probability

$$T(\sigma|\sigma') := \prod_{a=1}^2 T_a(\sigma_a|\sigma'). \quad (52)$$

By introducing a vector $\mathbf{P}^* := (P^*(\sigma))_{\sigma \in A^2}$ and a matrix $T' := T - \mathbb{I}_{M^2}$, where \mathbb{I}_k is a $k \times k$

identity matrix, this condition can be rewritten as

$$T' \mathbf{P}^* = \mathbf{0}. \quad (53)$$

Therefore, for a non-trivial solution \mathbf{P}^* to exist, $\det T' = 0$ must hold. On the other hand, because of the relation between a matrix T' and its adjugate matrix $\text{Adj}(T')$, we obtain⁶⁾

$$T' \text{Adj}(T') = (\det T') \mathbb{I}_{M^2} = \mathbf{O}_{M^2}, \quad (54)$$

where \mathbf{O}_k is a $k \times k$ zero matrix. Then, we find that \mathbf{P}^* is proportional to the every column of $\text{Adj}(T')$. By choosing the last column of $\text{Adj}(T')$ as \mathbf{P}^* , we obtain

$$\mathbf{P}^*(\boldsymbol{\sigma}) = C \cdot (\text{Adj}(T'))_{\sigma, (M, M)} (\forall \boldsymbol{\sigma}), \quad (55)$$

where C is a constant. By using this fact, the expected value of a quantity B with respect to the stationary distribution is

$$\begin{aligned} \langle B \rangle^* &= \sum_{\boldsymbol{\sigma}} B(\boldsymbol{\sigma}) \mathbf{P}^*(\boldsymbol{\sigma}) \\ &= CD(\mathbf{B}), \end{aligned} \quad (56)$$

where we have defined

$$D(\mathbf{B}) := \begin{vmatrix} \mathbf{T}'(1, 1)^\top \\ \mathbf{T}'(1, 2)^\top \\ \vdots \\ \mathbf{T}'(M, M-1)^\top \\ \mathbf{B}^\top \end{vmatrix}, \quad (57)$$

and vectors $\mathbf{T}'(\boldsymbol{\sigma}) := (T'(\boldsymbol{\sigma}|\boldsymbol{\sigma}'))_{\boldsymbol{\sigma}' \in A^2}$ and $\mathbf{B} := (B(\boldsymbol{\sigma}))_{\boldsymbol{\sigma} \in A^2}$. We find that $C = D(\mathbf{1})^{-1}$, where $\mathbf{1}$ is a vector of all ones. Then, a linear relation (50) can be rewritten as

$$0 = \frac{D(\alpha_1 \mathbf{s}_1 + \alpha_2 \mathbf{s}_2 + \alpha_0 \mathbf{1})}{D(\mathbf{1})}. \quad (58)$$

Below we set $\mathbf{B} = \alpha_1 \mathbf{s}_1 + \alpha_2 \mathbf{s}_2 + \alpha_0 \mathbf{1}$.

The necessary and sufficient condition for Eq. (58) to hold is that M^2 vectors $\mathbf{T}'(1, 1), \dots, \mathbf{T}'(M, M-1), \mathbf{B}$ are linearly dependent, that is

$$\sum_{\boldsymbol{\sigma} \neq (M, M)} c_{\boldsymbol{\sigma}} \mathbf{T}'(\boldsymbol{\sigma}) + c_{(M, M)} \mathbf{B} = \mathbf{0} \quad (59)$$

for some non-trivial $\{c_{\boldsymbol{\sigma}}\}$ (that is, not $c_{\boldsymbol{\sigma}} = 0 (\forall \boldsymbol{\sigma})$). When player 1 uses TFT, Eq. (59) is written as

$$0 = \sum_{\boldsymbol{\sigma} \neq (M, M)} c_{\boldsymbol{\sigma}} \left[\delta_{\sigma_1, \sigma'_2} T_2(\sigma_2|\boldsymbol{\sigma}') - \delta_{\sigma_1, \sigma'_1} \delta_{\sigma_2, \sigma'_2} \right] + c_{(M, M)} B(\boldsymbol{\sigma}') (\forall \boldsymbol{\sigma}'). \quad (60)$$

These are M^2 simultaneous equations of M^2 variables $\{c_\sigma\}$. Or, explicitly, the vectors $\mathbf{T}'(\sigma_1, \sigma_2)$ are

$$\mathbf{T}'(\sigma_1, \sigma_2) = \begin{pmatrix} 0 \\ \vdots \\ T_2(\sigma_2|1, \sigma_1) \\ \vdots \\ 0 \\ 0 \\ \vdots \\ T_2(\sigma_2|2, \sigma_1) \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ T_2(\sigma_2|M, \sigma_1) \\ \vdots \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \leftarrow (\sigma_1, \sigma_2) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (\forall \sigma \neq (M, M)), \quad (61)$$

and therefore Eq. (59) can be expressed as

$$\begin{pmatrix} \sum_{\sigma_2} c_{(1, \sigma_2)} T_2(\sigma_2|1, 1) \\ \vdots \\ \sum_{\sigma_2} c_{(M-1, \sigma_2)} T_2(\sigma_2|1, M-1) \\ \sum_{\sigma_2 \neq M} c_{(M, \sigma_2)} T_2(\sigma_2|1, M) \\ \vdots \\ \sum_{\sigma_2} c_{(1, \sigma_2)} T_2(\sigma_2|M, 1) \\ \vdots \\ \sum_{\sigma_2} c_{(M-1, \sigma_2)} T_2(\sigma_2|M, M-1) \\ \sum_{\sigma_2 \neq M} c_{(M, \sigma_2)} T_2(\sigma_2|M, M) \end{pmatrix} - \begin{pmatrix} c_{(1,1)} \\ \vdots \\ c_{(1, M-1)} \\ c_{(1, M)} \\ \vdots \\ c_{(M,1)} \\ \vdots \\ c_{(M, M-1)} \\ 0 \end{pmatrix} + c_{(M, M)} \mathbf{B} = \mathbf{0}. \quad (62)$$

We remark that the normalization condition of T_2 leads to

$$T_2(M|\sigma'_1, \sigma'_2) = 1 - \sum_{\sigma_2=1}^{M-1} T_2(\sigma_2|\sigma'_1, \sigma'_2) \quad (\forall \sigma'). \quad (63)$$

By using this fact, Eq. (62) is rewritten as

$$\begin{pmatrix} \sum_{\sigma_2 \neq M} (c_{(1,\sigma_2)} - c_{(1,M)}) T_2(\sigma_2|1, 1) \\ \vdots \\ \sum_{\sigma_2 \neq M} (c_{(M-1,\sigma_2)} - c_{(M-1,M)}) T_2(\sigma_2|1, M-1) \\ \sum_{\sigma_2 \neq M} c_{(M,\sigma_2)} T_2(\sigma_2|1, M) \\ \vdots \\ \sum_{\sigma_2 \neq M} (c_{(1,\sigma_2)} - c_{(1,M)}) T_2(\sigma_2|M, 1) \\ \vdots \\ \sum_{\sigma_2 \neq M} (c_{(M-1,\sigma_2)} - c_{(M-1,M)}) T_2(\sigma_2|M, M-1) \\ \sum_{\sigma_2 \neq M} c_{(M,\sigma_2)} T_2(\sigma_2|M, M) \end{pmatrix} + \begin{pmatrix} c_{(1,M)} \\ \vdots \\ c_{(M-1,M)} \\ 0 \\ \vdots \\ c_{(1,M)} \\ \vdots \\ c_{(M-1,M)} \\ 0 \end{pmatrix} - \begin{pmatrix} c_{(1,1)} \\ \vdots \\ c_{(1,M-1)} \\ c_{(1,M)} \\ \vdots \\ c_{(M,1)} \\ \vdots \\ c_{(M,M-1)} \\ 0 \end{pmatrix} + c_{(M,M)} \mathbf{B} = \mathbf{0}. \quad (64)$$

Since we consider the situation that TFT unilaterally enforces a linear relation between expected payoffs, this equation must hold irrespective of the strategy T_2 of player 2. Therefore, the first vector must be zero. We remark that the coefficient of $T_2(\sigma_2|i, j)$ in these equations is common for all i . This leads to

$$c_{(i,\sigma_2)} = c_i \quad (1 \leq i \leq M-1, 1 \leq \sigma_2 \leq M) \quad (65)$$

and

$$c_{(M,\sigma_2)} = 0 \quad (1 \leq \sigma_2 \leq M-1). \quad (66)$$

Below, we write $c_M := 0$. Substituting Eqs. (65) and (66) into Eq. (60), we finally obtain for $\forall \sigma'$

$$\begin{aligned} 0 &= \left(\sum_{\sigma}^{\sigma_1 \neq M} + \sum_{\sigma}^{\sigma_1 = M, \sigma_2 \neq M} \right) c_{\sigma} \left[\delta_{\sigma_1, \sigma_2} T_2(\sigma_2|\sigma') - \delta_{\sigma_1, \sigma_1'} \delta_{\sigma_2, \sigma_2'} \right] + c_{(M,M)} B(\sigma') \\ &= \left(\sum_{\sigma}^{\sigma_1 \neq M} + \sum_{\sigma}^{\sigma_1 = M, \sigma_2 \neq M} \right) c_{\sigma_1} \left[\delta_{\sigma_1, \sigma_2} T_2(\sigma_2|\sigma') - \delta_{\sigma_1, \sigma_1'} \delta_{\sigma_2, \sigma_2'} \right] + c_{(M,M)} B(\sigma') \\ &= \sum_{\sigma_1 \neq M} c_{\sigma_1} \delta_{\sigma_1, \sigma_2'} - \sum_{\sigma_1 \neq M} c_{\sigma_1} \delta_{\sigma_1, \sigma_1'} + c_{(M,M)} B(\sigma') \\ &= \sum_{\sigma_1} c_{\sigma_1} \delta_{\sigma_1, \sigma_2'} - \sum_{\sigma_1} c_{\sigma_1} \delta_{\sigma_1, \sigma_1'} + c_{(M,M)} B(\sigma') \\ &= c_{\sigma_2'} - c_{\sigma_1'} + c_{(M,M)} B(\sigma') \\ &= \sum_{\sigma_1} c_{\sigma_1} \hat{T}_1(\sigma_1|\sigma') + c_{(M,M)} B(\sigma'). \end{aligned} \quad (67)$$

Therefore, TFT is a ZD strategy. \square

Theorem 2 states that TFT cannot unilaterally enforce any linear relations between expected payoffs in non-potential games, if the opponent also uses memory-one strategies and a stationary distribution exists. Extension of this theorem to the case that the opponent uses memory- n strategies is a subject of future work.

5. Example

In this section, we provide two examples of potential game where TFT is a ZD strategy.

5.1 Two-player three-action game

We first consider the following two-player three-action symmetric zero-sum game:

$$\mathbf{s}_1 = (0, -2, -1, 2, 0, 1, 1, -1, 0)^\top \quad (68)$$

$$\mathbf{s}_2 = (0, 2, 1, -2, 0, -1, -1, 1, 0)^\top. \quad (69)$$

We can easily check that this game is a potential game with $c_1 = 1/2$, $c_2 = -3/2$ and $c_3 = -1/2$ in Lemma 2, and the potential is $\Phi = (0, 2, 1, 2, 4, 3, 1, 3, 2)^\top$ if we assume that $\Phi(1, 1) = 0$. We write the strategy of player a by $\mathbf{T}_a(\sigma) := (T_a(\sigma|\sigma))_{\sigma \in A^2}$. TFT of player 1 is $\mathbf{T}_1(1) = (1, 0, 0, 1, 0, 0, 1, 0, 0)^\top$, $\mathbf{T}_1(2) = (0, 1, 0, 0, 1, 0, 0, 1, 0)^\top$, and $\mathbf{T}_1(3) = (0, 0, 1, 0, 0, 1, 0, 0, 1)^\top$. By writing the Press-Dyson vectors as $\hat{\mathbf{T}}_a(\sigma) := (\hat{T}_a(\sigma|\sigma))_{\sigma \in A^2}$, we obtain $\hat{\mathbf{T}}_1(1) = (0, -1, -1, 1, 0, 0, 1, 0, 0)^\top$, $\hat{\mathbf{T}}_1(2) = (0, 1, 0, -1, 0, -1, 0, 1, 0)^\top$, and $\hat{\mathbf{T}}_1(3) = (0, 0, 1, 0, 0, 1, -1, -1, 0)^\top$. We can check that the relation

$$\sum_{\sigma=1}^3 c_\sigma \hat{\mathbf{T}}_1(\sigma) = \frac{1}{2} [\mathbf{s}_1 - \mathbf{s}_2] \quad (70)$$

indeed holds, which means that TFT is a ZD strategy.

When player 2 uses the memory-one strategy $\mathbf{T}_2(1) = (0, 1, 0, 0, 1, 0, 0, 1, 0)^\top$, $\mathbf{T}_2(2) = (0, 0, 1, 0, 0, 1, 0, 0, 1)^\top$, and $\mathbf{T}_2(3) = (1, 0, 0, 1, 0, 0, 1, 0, 0)^\top$, and both players choose the actions in the first round by the uniform probability distribution $(1/3, 1/3, 1/3)$, the time evolu-

tion is described by the Markov chain (49) with the transition probability

$$T = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (71)$$

and the initial condition $(1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9)^\top$. We can easily check that for this initial condition, the Markov chain converges to the stationary distribution $\mathbf{P}^* = (0, 0, 1/3, 1/3, 0, 0, 0, 1/3, 0)^\top$ at the second round. The expected payoffs are

$$\langle s_1 \rangle^* = \langle s_2 \rangle^* = 0, \quad (72)$$

which is consistent with Corollary 1.

When we consider a slightly different game¹²⁾

$$\mathbf{s}_1 = (0, 0, -1, 0, 0, 1, 1, -1, 0)^\top \quad (73)$$

$$\mathbf{s}_2 = (0, 0, 1, 0, 0, -1, -1, 1, 0)^\top, \quad (74)$$

this game is not a potential game. When the strategies of both players are the same as those above, we obtain

$$\langle s_1 \rangle^* = -\frac{2}{3} \quad (75)$$

$$\langle s_2 \rangle^* = \frac{2}{3}. \quad (76)$$

Therefore, player 1 cannot unilaterally enforce a linear relation $\langle s_1 \rangle^* = \langle s_2 \rangle^*$. In addition, because this game is also a zero-sum game, $\mathbf{s}_2 = -\mathbf{s}_1$ holds. Furthermore, it should be noted that $\hat{\mathbf{T}}_1(3) = -\hat{\mathbf{T}}_1(1) - \hat{\mathbf{T}}_1(2)$ holds due to the normalization condition of probability. Therefore, if TFT is a ZD strategy for this game, the relation

$$\sum_{\sigma=1}^2 d_\sigma \hat{\mathbf{T}}_1(\sigma) = \alpha_1 \mathbf{s}_1 + \alpha_0 \mathbf{1} \quad (77)$$

must hold with some non-trivial coefficients. Since the (1, 1) component of the left-hand side is zero, α_0 must be zero. Moreover, since the (1, 3) component of $\hat{\mathbf{T}}_1(2)$ and the (3, 2)

component of $\hat{T}_1(1)$ are zero, $d_1 = \alpha_1$ and $d_2 = -\alpha_1$ must hold. Then the (1, 2) component of the left-hand side is $-2\alpha_1$ and that of the right-hand side is zero, leading to contradiction. Therefore, TFT is not a ZD strategy in this game.

5.2 Cournot duopoly game

As noted in Section 3, the properties of ZD strategies and potential games hold even if the action set is uncountable. Moreover, the theoretical results in subsection 4.1 can also be easily extended to the case that the action is a continuous variable. Therefore, we here consider the Cournot duopoly game with unbounded payoffs. The action space of both players is $A = [0, \infty)$. The payoff of player a is given by

$$s_a(\boldsymbol{\sigma}) = \left[A - B \sum_{b=1}^2 \sigma_b \right] \sigma_a - C\sigma_a. \quad (78)$$

The Cournot duopoly game has a potential

$$\Phi(\boldsymbol{\sigma}) = \left[(A - C) - B \sum_{b=1}^2 \sigma_b \right] \sum_{b=1}^2 \sigma_b + B\sigma_1\sigma_2. \quad (79)$$

For this continuous action space, the definition (15) of a potential leads to

$$\frac{\partial s_a}{\partial \sigma_a}(\boldsymbol{\sigma}) = \frac{\partial \Phi}{\partial \sigma_a}(\boldsymbol{\sigma}) \quad (\forall \boldsymbol{\sigma}) \quad (80)$$

for all a . We can easily check that this relation indeed holds for Eqs. (78) and (79).

If player a takes TFT

$$T_a(\sigma_a|\boldsymbol{\sigma}') = \delta(\sigma_a - \sigma'_{-a}), \quad (81)$$

where δ represents the Dirac delta function, her Press-Dyson vectors are

$$\hat{T}_a(\sigma_a|\boldsymbol{\sigma}') = \delta(\sigma_a - \sigma'_{-a}) - \delta(\sigma_a - \sigma'_a). \quad (82)$$

When we consider the quantity

$$c(\boldsymbol{\sigma}) = [(A - C) - B\boldsymbol{\sigma}] \boldsymbol{\sigma}, \quad (83)$$

we obtain

$$\begin{aligned} \int d\boldsymbol{\sigma} c(\boldsymbol{\sigma}) \hat{T}_a(\boldsymbol{\sigma}|\boldsymbol{\sigma}') &= c(\boldsymbol{\sigma}'_{-a}) - c(\boldsymbol{\sigma}'_a) \\ &= [(A - C) - B\boldsymbol{\sigma}'_{-a}] \boldsymbol{\sigma}'_{-a} - [(A - C) - B\boldsymbol{\sigma}'_a] \boldsymbol{\sigma}'_a \\ &= \left[(A - C) - B \sum_{b=1}^2 \sigma'_b \right] \boldsymbol{\sigma}'_{-a} - \left[(A - C) - B \sum_{b=1}^2 \sigma'_b \right] \boldsymbol{\sigma}'_a \\ &= s_{-a}(\boldsymbol{\sigma}') - s_a(\boldsymbol{\sigma}'). \end{aligned} \quad (84)$$

Therefore, TFT is a ZD strategy, and unilaterally enforces $\langle s_1 \rangle^* = \langle s_2 \rangle^*$.

6. Discussion

In Ref.,¹⁴⁾ the authors proved that TFT is unbeatable if and only if the stage game is a potential game. In this section, we introduce their results and discuss the relation between our results and their results.

We first introduce the unbeatable property.¹⁴⁾

Definition 5 *The strategy of player a is unbeatable if*

$$\sum_{t=1}^T [s_{-a}(\sigma^{(t)}) - s_a(\sigma^{(t)})] \leq \max_{\sigma} [s_{-a}(\sigma) - s_a(\sigma)] \quad (\forall T \geq 1) \quad (85)$$

for any strategies $\{T_{-a}^{(t)}(\sigma_{-a}^{(t)}|h_{[1:t-1]})\}_{t=1}^{\infty}$ of player $-a$.

Duersch et al. proved the following proposition.

Proposition 2 (¹⁴⁾) *For two-player symmetric games, TFT is unbeatable if and only if the stage game G is a potential game.*

When combined with our results, the following three conditions are equivalent for infinitely repeated two-player symmetric games: (i) The stage game is a potential game, (ii) TFT is unbeatable, (iii) TFT is a ZD strategy which unilaterally enforces $\mathcal{S}_1 = \mathcal{S}_2$. Particularly, TFT is unbeatable if and only if TFT is a ZD strategy.

A natural question is whether the equivalence of unbeatable property and a ZD strategy also holds for other strategies including imitation strategies in potential games. Generally, unbeatable property is easily interpreted and observed in other strategies,¹²⁾ but extension to other games may be difficult, since difference of payoffs of two players is not always important for other games. On the other hand, the concept of ZD strategy is clearly defined in general games. Clearly, ZD strategies are not necessarily unbeatable, as we can see for the case of an equalizer strategy.⁶⁾ Meanwhile, here we consider the *imitate-if-better* strategy of player a ¹²⁾

$$T_a(\sigma_a|\sigma') = \delta_{\sigma_a, \sigma'_a} \mathbb{I}(s_{-a}(\sigma') > s_a(\sigma')) + \delta_{\sigma_a, \sigma'_a} \mathbb{I}(s_{-a}(\sigma') \leq s_a(\sigma')) \quad (\forall \sigma_a, \forall \sigma'), \quad (86)$$

where $\mathbb{I}(\dots)$ is an indicator function that returns 1 when \dots holds and 0 otherwise. This strategy has been known to be unbeatable in potential games.¹⁴⁾ By using Lemma 2, we find that

$$\sum_{\sigma_a} c_{\sigma_a} \hat{T}_a(\sigma_a|\sigma') = \sum_{\sigma_a} c_{\sigma_a} \{ \delta_{\sigma_a, \sigma'_a} - \delta_{\sigma_a, \sigma'_a} \} \mathbb{I}(s_{-a}(\sigma') > s_a(\sigma'))$$

$$= \left\{ \frac{1}{2} s_a(\sigma') - \frac{1}{2} s_{-a}(\sigma') \right\} \mathbb{I}(s_{-a}(\sigma') > s_a(\sigma')) \quad (87)$$

for potential games, where the function c_σ is that in Eq. (20). Then, from Lemma 1, we obtain the following proposition.

Proposition 3 *For two-player symmetric games, the imitate-if-better strategy (86) of player a unilaterally enforces*

$$0 = \langle \{s_a(\sigma') - s_{-a}(\sigma')\} \mathbb{I}(s_{-a}(\sigma') > s_a(\sigma')) \rangle^* \quad (88)$$

for potential games.

This is not a ZD strategy, but is contained in the class of extended ZD strategies which unilaterally enforce linear relations between conditional expectations of payoffs.²⁵⁾ In fact, Eq. (88) implies that the probability that the state σ' such that $s_{-a}(\sigma') > s_a(\sigma')$ is realized is zero. Therefore, although unbeatable strategies are not necessarily ZD strategies, they may be contained in the class of extended ZD strategies. Further investigation is needed to this topic.

Finally, we discuss a slight difference between Ref.¹⁴⁾ and our results. We first introduce the following concept.

Definition 6 *The strategy of player a is weakly unbeatable if*

$$\langle s_a \rangle^* \geq \langle s_{-a} \rangle^* \quad (89)$$

for any strategies of player $-a$.

It should be noted that an unbeatable strategy is weakly unbeatable. In Ref.,¹⁷⁾ weakly unbeatable strategies are called *competitive* (or *rival*) strategies. The following corollary is a direct consequence of Proposition 2.

Corollary 2 *For two-player symmetric games, TFT is weakly unbeatable for potential games. In other words, when player a uses TFT, then*

$$\langle s_a \rangle^* \geq \langle s_{-a} \rangle^*. \quad (90)$$

This corollary is weaker than Corollary 1, because Corollary 1 claims that the equality $\langle s_a \rangle^* = \langle s_{-a} \rangle^*$ must hold in the inequality (90). In other words, although TFT is unbeatable, TFT can also never win. On the other hand, as we can see from Eq. (88), the imitate-if-better strategy can win. For example, in the prisoner's dilemma game, when player $-a$ uses the strategy which always cooperates (All-C), the imitate-if-better strategy wins if she starts with defection.

7. Concluding Remarks

In this paper, we proved that, for infinitely repeated two-player symmetric games, TFT is a ZD strategy, which unilaterally enforces $\langle s_1 \rangle^* = \langle s_2 \rangle^*$, if and only if the stage game is a potential game. We also proved that, TFT cannot unilaterally enforce any linear relations between expected payoffs in non-potential games, if the opponent also uses memory-one strategies and a stationary distribution of the induced Markov chain exists. We explicitly showed that TFT is a ZD strategy in the two potential games, that is, a two-player three-action zero-sum game and the Cournot duopoly game. Furthermore, we proved that the imitate-if-better strategy can be regarded as an extended ZD strategy in potential games. When combined with the results of Duersch et al.,¹⁴⁾ which proved that TFT is unbeatable if and only if the stage game is a potential game, TFT is unbeatable if and only if TFT is a ZD strategy. This result suggests that there may be some relations between unbeatable strategies and ZD strategies.

In this paper, we consider only two-player symmetric games. Extension of our result to multi-player symmetric games is non-trivial, because TFT in multi-player symmetric games cannot be defined uniquely, although there are many multi-player symmetric potential games. We would like to investigate whether imitation strategies in multi-player symmetric games are ZD strategies or not in future.

Another subject of future work is whether TFT is efficient in asymmetric games. Even though the game is asymmetric, imitation would be useful if the payoff of a player is similar to that of other player. We would like to find the condition in which TFT is useful in asymmetric games.

Acknowledgment

This study was supported by JSPS KAKENHI Grant Number JP20K19884.

References

- 1) A. Rapoport, A. M. Chammah, and C. J. Orwant: *Prisoner's dilemma: A study in conflict and cooperation* (University of Michigan press, 1965), Vol. 165.
- 2) R. Axelrod and W. D. Hamilton: *Science* **211** (1981) 1390.
- 3) M. A. Nowak and K. Sigmund: *Nature* **355** (1992) 250.
- 4) M. Nowak and K. Sigmund: *Nature* **364** (1993) 56.
- 5) L. A. Imhof, D. Fudenberg, and M. A. Nowak: *Journal of Theoretical Biology* **247** (2007) 574.
- 6) W. H. Press and F. J. Dyson: *Proceedings of the National Academy of Sciences* **109** (2012) 10409.
- 7) M. Ueda: *Journal of the Physical Society of Japan* **90** (2021) 025002.
- 8) S. D. Yi, S. K. Baek, and J.-K. Choi: *Journal of Theoretical Biology* **412** (2017) 1.
- 9) Y. Murase and S. K. Baek: *Scientific Reports* **10** (2020) 16904.
- 10) F. Vega-Redondo: *Econometrica: Journal of the Econometric Society* **65** (1997) 375.
- 11) K. H. Schlag: *Journal of Economic Theory* **78** (1998) 130.
- 12) P. Duersch, J. Oechssler, and B. C. Schipper: *Games and Economic Behavior* **76** (2012) 88.
- 13) L. C. Baiardi and A. K. Naimzada: *Applied Mathematics and Computation* **336** (2018) 193.
- 14) P. Duersch, J. Oechssler, and B. C. Schipper: *International Journal of Game Theory* **43** (2014) 25.
- 15) D. Monderer and L. S. Shapley: *Games and Economic Behavior* **14** (1996) 124.
- 16) D. Hao, Z. Rong, and T. Zhou: *Phys. Rev. E* **91** (2015) 052803.
- 17) C. Hilbe, A. Traulsen, and K. Sigmund: *Games and Economic Behavior* **92** (2015) 41.
- 18) X. He, H. Dai, P. Ning, and R. Dutta: *IEEE Signal Processing Letters* **23** (2016) 311.
- 19) A. McAvoy and C. Hauert: *Theoretical Population Biology* **113** (2017) 13.
- 20) M. Ueda: *Royal Society Open Science* **8** (2021) 202186.
- 21) D. Fudenberg and J. Tirole: *Game Theory* (MIT Press, Massachusetts, 1991).
- 22) E. Akin: *Ergodic Theory, Advances in Dynamical Systems* (2016) 77.
- 23) M. Ueda and T. Tanaka: *PLOS ONE* **15** (2020) e0230973.

- 24) A. McAvoy and C. Hauert: *Proceedings of the National Academy of Sciences* **113** (2016) 3573.
- 25) M. Ueda: arXiv preprint arXiv:2012.10231 (2020).
- 26) P. Duersch, J. Oechssler, and B. C. Schipper: *International Journal of Game Theory* **41** (2012) 553.
- 27) A. Mamiya and G. Ichinose: *Journal of Theoretical Biology* **477** (2019) 63.