Nonlinear Functional Analysis and Applications
Vol. 31, No. 2 (2026), pp. 317-342
ISSN: 1229-1595(print), 2466-0973(online)

https://doi.org/10.22771/nfaa.2026.31.02.01
http://nfaa.kyungnam.ac.kr/journal-nfaa

ZERO-DETERMINANT STRATEGIES OF THREE-PLAYER REPEATED GAMES UNDER NOISY OBSERVATION

S. M. Abd El-Salam[1], E. El-Seidy[2], A. R. Abdel-Malek[3] and A. Rabie[4]

[1]Department of Mathematics, Faculty of Science, Ain Shams University, Cairo, Egypt
e-mail: `Salsabeel.mohamed@cis.asu.edu.eg`

[2]Department of Mathematics, Faculty of Science, Ain Shams University, Cairo, Egypt
e-mail: `esam_elsedy@hotmail.com`

[3]Department of Mathematical and Natural Sciences, Faculty of Engineering, Egyptian Russian University, Badr, Egypt
e-mail: `amira-ragab@eru.edu.eg`

[4]Planning Techniques Center, Institute of National Planning, Cairo, Egypt
e-mail: `aya.ibrahim@inp.edu.eg`

Abstract. The emergence of zero-determinant (ZD) strategies has generated considerable interest, particularly in the context of the Iterated Prisoner's Dilemma (IPD). In this game, players sustain cooperation to avoid retaliation in subsequent rounds. This study investigates the application of ZD strategies within a three-player repeated game framework, with a particular focus on the influence of observation errors. We analyze how ZD strategies interact with the inherent incompleteness of observations, demonstrating that even under imperfect monitoring, certain strategies can still enforce linear payoff relationships. Both the discount factor and observation errors are considered in the analysis, and a numerical example is provided to illustrate the robustness and limitations of ZD strategies in a non-3×3 Prisoner's Dilemma game. These results contribute to a deeper understanding of strategic decision-making in repeated games and highlight the potential of ZD strategies to control long-term outcomes even amidst observational uncertainty.

---

[0]Received August 14, 2025. Revised December 9, 2025. Accepted December 13, 2025.
[0]2020 Mathematics Subject Classification: 46E35, 49K20, 76D05, 35Q30.
[0]Keywords: Zero-determinant strategies, iterated games, symmetric games, prisoner's dilemma game, discounting factor, observational uncertainties.
[0]Corresponding author: A. Rabie(`aya.ibrahim@inp.edu.eg`).

1. Introduction

The dilemma game is a central concept in evolutionary and classical game theory, commonly known as the PD game, with wide applications in areas such as political science and environmental studies [32]. Cooperation is essential for sustainable societies, yet it is often costly for individuals, reducing the likelihood of cooperative behavior. In game theory, "Perfect Monitoring" means that players have complete and accurate information about opponents' actions and payoffs. In the iterated prisoner's dilemma (IPD), perfect monitoring implies that cooperation "C" always produces signal "g", and defection "D" produces signal "b", allowing the game to model behavior across repeated rounds [2].

Under perfect monitoring, players fully observe opponents' choices, eliminating uncertainty and enabling decisions based on actual past actions. Experiments in [3] examined how monitoring structures influence infinitely repeated PD dynamics, while [4] introduced new algorithms for computing equilibrium payoffs. Perfect monitoring thus simplifies strategic decision-making, enabling players to adjust behavior predictably across rounds. A method for deriving payoff pairs for pure strategy subgame-perfect equilibria with public randomization was presented in [5].

In reality, perfect monitoring is often unattainable, prompting the use of private or imperfect monitoring to reflect real-world repeated interactions better. Press and Dyson first discovered zero determinant strategies in 2012 marked a turning point in the study of the recurring (IPD). A player employing a ZD strategy can impose a linear relationship between their expected payoff and that of their opponents, regardless of the opponents' strategies. This allows the player to impose, equalize, or control long-term payoffs in recurring interactions. This discovery transformed the theoretical landscape of recurring games, demonstrating the possibility of strategic control even without cooperation from other players [6]. Analytical studies of ZD strategies under a 2×2 payoff matrix were presented in [7].

ZD strategies create fixed linear relationships between players' payoffs, making extortion strategies [8], [9] particularly notable. Even when player β has superior memory, the score of player α remains unaffected compared to a specific finite-memory strategy used by β, as shown in [6]. The discovery of ZD strategies stimulated extensive research [26]–[12], especially after Stewart and Plotkin's question in [13], expanding into multiplayer games [14], asymmetric games [15], continuous or alternating action spaces [16], animal contests [17], human experiments [18], and responses to computerized ZD strategies [19]. Updated classifications of allies ("excellent strategies") and opponents appear in [20], [21], with recent applications extending to technical fields and human interaction [22]. Effects on cooperation and conflict resolution in repeated games are further examined in [23].

Early ZD studies assumed no errors, although noise is inherent in human interactions and can disrupt cooperation. Many works addressed errors in IPD more broadly [24], [25], but only a few examined their impact on ZD strategies [27], [28]. Errors are generally categorized into implementation errors [29] and perception errors [30]. Hao [27] and Mamiya and Ichinose [28] considered perception-based errors arising from limited access to opponents' direct actions, forcing players to rely solely on private monitoring.

Queuing systems can be viewed as recurring interactions between servers and customers, where each decision such as joining, waiting, or providing service affects future system states. This dynamic nature allows the system to be modeled as a recurring game. Therefore, ZD strategies can be used to enable one player (typically the server) to unilaterally impose a linear relationship between its long-term performance metrics (such as service cost) and those of other players (such as wait time or queue length). ZD strategies thus provide an effective tool for designing service policies that control queue behavior and regulate customer responses.

This research shows that ZD strategies are still possible when such observation errors are present. These studies don't take into account the discount element. It makes sense to believe that rewards in the future will be discounted. Consequently, various research efforts [15] have focused on analyzing a discount factor applicable to ZD strategies and have mathematically derived the minimal discount factor required for the existence of ZD strategies [18]. Strategies with ZD possessing the ability to gain informational advantage, facilitating mutual identification and ultimately leading to the attainment of an evolutionarily stable state, are exhibited [4]. ZD strategies with a discount factor and observation errors for two players are studied in [23].

In this study, our objective is to investigate ZD strategies considering both observation errors and a discount factor. Moreover, we aim to identify alternative strategies that impose a linear correlation in the payoff of the three players. The determinants of the expected payoffs in the IPD game are mathematically defined.

The following sections outline the structure of the paper. In Section 2, In this part, you will find an explanation of the IPD game and the interest strategies, which are also known as memory-one strategies. Additionally, details about the anticipated payoffs are provided. In sections 3 and 4, we proceeded to build the payoff matrix for three players by incorporating observation errors and discount factors. Our aim was to explore various strategies in order to determine payoffs associated with varying levels of errors. Our focus in section 5 lies on the six strategies commonly referred to as S_{36}, S_{38}, S_{52}, S_{54}, S_{63} and S_0 respectively. Afterward, we determine the payoff values by analyzing the errors, which are then presented in a structured table format. Additionally, we create a visual representation illustrating the dominant strategies' behavior and the paper is concluded in Section 6.

2. ZD strategies in three players

In this section, the IPD game and interest strategies known as memory-one strategies are explained as well as the expected payoff. There is a comprehensive discussion of these strategies and the expected payoff. An analogous three-person PD game is assumed to be considered in which the reward matrix is presented.

The PD is a model of cooperation and competition between individuals who must make decisions about cooperation C (cooperation for the common good) or defect D (non-cooperation) in the face of others. Each player's reward depends on his and the other players' decisions in each round. The goal is to find the strategy that improves each player's expected reward in the long run.

Memory-one strategies in the context of game theory refer to strategies where players base their decisions solely on the outcomes of the previous rounds. The term "memory-one" arises from the fact that these strategies use only one memory to store and recall past results, simplifying the decision-making process. These strategies are particularly associated with their simplicity, making them easy to apply and analyze.

Theorem 2.1. *Suppose player X uses a strategy p with Press-Dyson vector*

$$\tilde{p} = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$$

and player Y employs a strategy that induces a sequence of distributions

$$\left\{ v^{(n)} \right\}_{n=1}^{\infty}.$$

Let v be any associated limit distribution. Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left(v^{(k)} \cdot \tilde{p} \right) = 0.$$

Therefore,

$$v \cdot \tilde{p} = v_1 p_{CC} + v_2 p_{CD} + v_3 p_{DC} + v_4 p_{DD} = 0.$$

Proof. Let the probability that player X cooperates in the n^{th} round be denoted by $v_{12}^{(n)}$. Then,

$$v_{12}^{(n)} = v_1^{(n)} + v_2^{(n)} = v^{(n)} \cdot e_{12},$$

where $e_{12} = (1, 1, 0, 0)^T$ is a selector vector that extracts the probability mass corresponding to the states where X cooperates.

As $n \to \infty$, if the limit distribution v exists, then the long-run average cooperation rate becomes:

$$\lim_{n\to\infty} v_{12}^{(n)} = v \cdot e_{12}.$$

This confirms that cooperation probabilities over time converge to a linear form dependent on the stationary distribution and the state classification. Since X is using the strategy p, p_i is the probability that X plays C in the next round, given the outcome of the current round, thus

$$v_{12}^{(n+1)} = v^{(n)} \cdot p.$$

The probability that X cooperates in the $(n+1)^{th}$ round. Hence,

$$v_{12}^{(n+1)} - v_{12}^{(n)} = v^{(n)} \cdot p - v^{(n)} \cdot e_{12} = v^{(n)} \left(p - e_{12}\right) = v^{(n)} \cdot \tilde{p}.$$

This implies,

$$v_{12}^{(n+1)} - v_{12}^{(1)} = \sum_{k=1}^{n} \left(v_{12}^{(k+1)} - v_{12}^{(k)}\right) = \sum_{k=1}^{n} \left(v^{(k)} \cdot \tilde{p}\right).$$

Since $0 \leq v_{12}^{(k)} \leq 1$ for any k, that is, the absolute value of the left side is at most 1,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left(v^{(k)} \cdot \tilde{p}\right) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left(v_{12}^{(k+1)} - v_{12}^{(k)}\right) = 0.$$

The sequence of probability vectors

$$\left\{ \frac{1}{n} \sum_{k=1}^{n} v_{12}^{(k)} \right\}$$

converges to the stationary distribution v, then the continuity of the dot product implies $v \cdot \tilde{p} = 0$. □

In the case of a three-player Prisoner's Dilemma (PD) game, each player's decision in a given round depends on the outcomes of the previous rounds.

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} CC & CD & DC & DD \end{array} \\ \begin{pmatrix} R & K & K & S \\ T & L & L & P \end{pmatrix} \end{array}. \tag{2.1}$$

The rewards accumulated by the pivot player are represented by the values within the input matrix. The stated values represent the rewards earned by the central player, denoted by "α", for one round of a three-player repetitive game. Each row and column of the matrix represents the actions of the central player "α", another of the players, whom we denote by "β", and another third player we denote by "γ", respectively. From [26], we have

$$T > R > L > K > P > S \quad \text{and} \quad 2R > T + S.$$

When α the main player encounters players β and γ in one round, the first row represents the actions of player α (such as cooperation C or defect D) and the first column represent the actions of players β and γ (also C or D). The item in the selected cell of a particular row and column then displays the payoff to the pivot player α when they choose their action in the row and players β and γ choose their action in the corresponding column.

In short, the matrix describes how to determine the rewards that the pivotal player α receives for choosing a particular action based on players β, γ' action in one round of the game.

Players α, β, and γ formulate their initial memory strategies by incorporating the result of the preceding round in order to make informed decisions in the current round. These strategies are represented by a set of values that represent the players' choices in the various possible situations.

Speaking of player α's first memory strategies, the strategy can be represented by an 8-tuple $(\jmath_1, \jmath_2, \jmath_3, \jmath_4, \jmath_5, \jmath_6, \jmath_7, \jmath_8)$ where:

- $\jmath_1$ is the probability that player α selects option C (cooperation), players β and γ follow,
- $\jmath_2$ is the probability that player α decides to choose C (cooperation), player β also chooses C, and player γ decides to choose D,
- $\jmath_3$:Once player α opts for C (cooperation), player β goes for option D, while player γ decides to follow α's choice by selecting option C as well,
- $\jmath_4$ is the probability that player α selects option C (cooperation), followed by player β and player γ selecting option D,
- $\jmath_5$ is the probability that player α chooses choice D (defect), player β chooses choice C, and player γ chooses C,
- $\jmath_6$ is the probability that player α chooses choice D (defect), player β chooses choice C, and player γ chooses D,
- $\jmath_7$ is the probability that player α chooses choice D (defect), player β chooses choice D, and player γ chooses C,
- $\jmath_8$ is the probability that player α chooses choice D (defect), player β chooses choice D, and player γ chooses D,

where $0 \leq \jmath_i \leq 1$, $i \in \{1, 2, \ldots, 8\}$.

The decisions and first memory methods of player α in various game scenarios with players β and γ may be identified in this way. According to the probabilities of player α and other players acting in different ways, the values of the options can change. The first memory strategies for the game can be expressed through the utilization of the payoff matrix.

In each round t, players make their move based on the outcome of the previous round (round $t-1$) and the other players' random odds. Thus, the given values of the previous choices can be used to determine the random state in the current round (round t), and so on. This process is repeated over repeated rounds of the game so that past decisions influence future decisions of the players.

$$\begin{aligned} V(t) = (&v_{CCC}(t),\ v_{CCD}(t),\ v_{CDC}(t),\ v_{CDD}(t), \\ &v_{DCC}(t),\ v_{DCD}(t),\ v_{DDC}(t),\ v_{DDD}(t)), \end{aligned} \tag{2.2}$$

where $v_{CCD}(t)$ represents the probability of α and β cooperating while γ chooses to defect in round t. This means that α and β choose the action C (cooperate) while γ chooses D (defect) in the same round. The normalization is given by

$$\begin{aligned} &v_{CCC}(t) + v_{CCD}(t) + v_{CDC}(t) + v_{CDD}(t) + v_{DCC}(t) \\ &+ v_{DCD}(t) + v_{DDC}(t) + v_{DDD}(t) = 1, \\ &\qquad t = 0, 1, \ldots . \end{aligned} \tag{2.3}$$

Remark 1. Assume $v = (v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8)$ is a unique stationary distribution. Then the expected payoffs to the players X, Y and Z, $\pi_X = v,\ S_X$, $\pi_Y = v,\ S_Y$; $\pi_Z = v,\ S_Z$, satisfy that $\pi_X \geq \pi_Y$ and $\pi_X \geq \pi_Z$ if and only if $v_5 \geq v_3$, $v_5 \geq v_2$ and $v_6 \geq v_4$, $v_7 \geq v_4$.

Proof. Since $\pi_X - \pi_Y = (v_5 - v_3)(T - K) + (v_6 - v_4)(L - S)$, then

- $\pi_X \geq \pi_Y$ if and only if $v_5 \geq v_3$ and $v_6 \geq v_4$,
- $\pi_Y - \pi_Z = (v_5 - v_2)(T - K) + (v_7 - v_4)(L - S)$,
- $\pi_X \geq \pi_Z$ if and only if $v_5 \geq v_2$ and $v_7 \geq v_4$,
- $\pi_X \geq \pi_Y \geq \pi_Z$ if and only if $v_5 \geq v_3 \geq v_2$ and $v_7 \geq v_6 \geq v_4$,
- $\pi_X = \pi_Y = \pi_Z$ if and only if $v_5 = v_3 = v_2$ and $v_7 = v_6 = v_4$.

□

The initial condition is determined by

$$\begin{aligned} V(0) = [&\jmath_0\kappa_0 z_0, \jmath_0\kappa_0(1-z_0), \jmath_0(1-\kappa_0)z_0, \jmath_0(1-\kappa_0)(1-z_0), (1-\jmath_0)\kappa_0 z_0, \\ &(1-\jmath_0)\kappa_0(1-z_0), (1-\jmath_0)(1-\kappa_0)z_0, (1-\jmath_0)(1-\kappa_0)(1-z_0)], \end{aligned} \tag{2.4}$$

where $\jmath_0$ is the probability that α cooperates in the first round, κ_0 is the probability that β cooperates in the first round and z_0 is the probability that γ cooperates in the first round.

The outcome of player α is $V(t)N_\alpha^T$ in round t, where

$$N_\alpha = (R, K, K, S, T, L, L, P). \tag{2.5}$$

In the repeated game, player α can anticipate a per-round payoff equivalent to

$$n_\alpha = (1-\delta)\sum_{t=0}^{\infty} \delta^t V(t) N_\alpha,\ 0 < \delta \leq 1. \tag{2.6}$$

From matrix (2.1) Which determines the rewards collected by the payoff of the player based on the actions of the three players. In relation (2.6), the parameter δ represents the discount factor of the repeated game. It measures how future payoffs are valued relative to the present ones, with $0 < \delta \leq 1$. A higher δ indicates that players place greater weight on future interactions, while a lower δ means that future payoffs are discounted more strongly.

The matrix of transition probability is

$$\text{A}=\left[\begin{matrix}
\jmath_1\kappa_1 z_1 & \jmath_1\kappa_1(1-z_1) & \jmath_1(1-\kappa_1)z_1 & \jmath_1(1-\kappa_1)(1-z_1) & (1-\jmath_1)\kappa_1 z_1 \\
\jmath_2\kappa_2 z_5 & \jmath_2\kappa_2(1-z_5) & \jmath_2(1-\kappa_2)z_5 & \jmath_2(1-\kappa_2)(1-z_5) & (1-\jmath_2)\kappa_2 z_5 \\
\jmath_3\kappa_5 z_2 & \jmath_3\kappa_5(1-z_2) & \jmath_3(1-\kappa_5)z_2 & \jmath_3(1-\kappa_5)(1-z_2) & (1-\jmath_3)\kappa_5 z_2 \\
\jmath_4\kappa_6 z_6 & \jmath_4\kappa_6(1-z_6) & \jmath_4(1-\kappa_6)z_6 & \jmath_4(1-\kappa_6)(1-z_6) & (1-\jmath_4)\kappa_6 z_6 \\
\jmath_5\kappa_3 z_2 & \jmath_5\kappa_3(1-z_3) & \jmath_5(1-\kappa_3)z_3 & \jmath_5(1-\kappa_3)(1-z_3) & (1-\jmath_5)\kappa_3 z_3 \\
\jmath_6\kappa_4 z_7 & \jmath_6\kappa_4(1-z_7) & \jmath_6(1-\kappa_4)z_7 & \jmath_6(1-\kappa_4)(1-z_7) & (1-\jmath_6)\kappa_4 z_6 \\
\jmath_7\kappa_7 z_4 & \jmath_7\kappa_7(1-z_4) & \jmath_7(1-\kappa_7)z_4 & \jmath_7(1-\kappa_7)(1-z_4) & (1-\jmath_7)\kappa_7 z_4 \\
\jmath_8\kappa_8 z_8 & \jmath_8\kappa_8(1-z_8) & \jmath_8(1-\kappa_8)z_8 & \jmath_8(1-\kappa_8)(1-z_8) & (1-\jmath_8)\kappa_8 z_8
\end{matrix}\right.$$

$$\left.\begin{matrix}
(1-\jmath_1)\kappa_1(1-z_1) & (1-\jmath_1)(1-\kappa_1)z_1 & (1-\jmath_1)(1-\kappa_1)(1-z_1) \\
(1-\jmath_2)\kappa_2(1-z_5) & (1-\jmath_2)(1-\kappa_2)z_5 & (1-\jmath_2)(1-\kappa_2)(1-z_5) \\
(1-\jmath_3)\kappa_5(1-z_2) & (1-\jmath_3)(1-\kappa_5)z_2 & (1-\jmath_3)(1-\kappa_5)(1-z_2) \\
(1-\jmath_4)\kappa_6(1-z_6) & (1-\jmath_4)(1-\kappa_6)z_6 & (1-\jmath_4)(1-\kappa_6)(1-z_6) \\
(1-\jmath_5)\kappa_3(1-z_3) & (1-\jmath_5)(1-\kappa_3)z_3 & (1-\jmath_5)(1-\kappa_3)(1-z_3) \\
(1-\jmath_6)\kappa_4(1-z_6) & (1-\jmath_6)(1-\kappa_4)z_7 & (1-\jmath_6)(1-\kappa_{C4})(1-z_7) \\
(1-\jmath_7)\kappa_7(1-z_4) & (1-\jmath_7)(1-\kappa_7)z_4 & (1-\jmath_7)(1-\kappa_7)(1-z_4) \\
(1-\jmath_8)\kappa_8(1-z_8) & (1-\jmath_8)(1-\kappa_8)z_8 & (1-\jmath_8)(1-\kappa_8)(1-z_8)
\end{matrix}\right]. \tag{2.7}$$

By substituting

$$V(t) = V(0)A \tag{2.8}$$

in Eq.(2.6), one gets

$$\begin{aligned} n_\alpha &= (1-\delta)V(0)\sum_{t=0}^{\infty}(\delta A)^t N_\alpha \\ &= (1-\delta)V(0)(I-\delta A)^{-1} N_\alpha, \end{aligned} \tag{2.9}$$

where I is the 8×8 identity matrix. Likewise, player β can anticipate a per-round payoff equivalent to

$$n_\beta = (1-\delta)V(0)\sum_{t=0}^{\infty}(\delta A)^t N_\beta, \tag{2.10}$$

where

$$N_\beta = (R, K, T, L, K, S, L, P). \tag{2.11}$$

Similarly

$$n_\gamma = (1-\delta)V(0)\sum_{t=0}^{\infty}(\delta A)^t N_\gamma, \tag{2.12}$$

where

$$N_\gamma = (R, T, K, L, K, L, S, P). \tag{2.13}$$

This lemma provides the necessary condition for deriving the stationary distribution, which we later use in Section 4 to prove the existence of ZD strategies under noise.

3. Model

We analyze the symmetric two-agent iterative prisoner's dilemma game characterized by private observation, drawing upon the existing scholarly literature.

3.1. **IPD with private monitoring.** The phrase IPD with private monitoring is employed in the field of game theory to depict a modified version of the typical PD game. In this rendition, the game is played iteratively within a limited number of rounds, and the participants possess confidential information or individual observations pertaining to the actions and consequences of the game.

In this game, three players are presented with the choice to either collaborate or deceive each other. The payoff matrix often represents the rewards for each participant depending on their actions together.

In a recurring version of the game (IPD), the game is played multiple times, allowing players to monitor each other's actions and rewards from previous rounds. Private monitoring means that each player has access to their previous actions and rewards but cannot see the other player's actions and rewards directly.

The concept of special observation is necessary in many real-world situations where players may have limited or incomplete information about their opponents' strategies and outcomes. It adds an element of uncertainty to the decision-making process and can lead to more complex strategies and behaviors.

In this particular game, every participant is notified about the other player's performance in the previous round. This notification is considered confidential information. Previous researches on repeated games involving private monitoring failed to examine the magnitude of effective equilibria's existence in relation to the IPD game [28].

In every round, each player $i \in \{\alpha, \beta, \gamma\}$ selects an action a_i from a set $\{C, D\}$, where C signifies cooperation and D represents defection. Once both players have completed their actions, player i assesses their own action a_i as well as the action of their opponent $x_i \in \{g, b\}$, with g representing a positive outcome and b representing a negative outcome. Because in the IPD with Private Observation, players are unable to fully monitor each other's actions. This makes collaboration more difficult, but it is still possible with strategies such as the TFT principle and the grim trigger. The potential for cooperation depends on the players' strategies, game standards, and willingness to trust each other.

$\phi(x|a)$ is the probability of receiving signal x given that you took action a. This probability depends on the signaling mechanism used in the game. You can use $\phi(x|a)$ to update your beliefs about your opponent's strategy.

The signal profile x is represented as $(x_\alpha, x_\beta, x_\gamma)$, where $x_\alpha, x_\beta, x_\gamma$ denote the signals that are received by players α, β, and γ, respectively. Each player receives a signal, and this signal profile represents the combination of all the signals that the three players receive in the game $a = (a_\alpha, a_\beta, a_\gamma)$. This is an action profile, where $a_\alpha, a_\beta, a_\gamma$ are the actions taken by players α, β, and γ, respectively. In a game, every player selects an action, and this action profile symbolizes the amalgamation of all the actions opted for by the three players [28].

Let ε be the probability of error for one player. It represents the likelihood that a player receives an incorrect signal, i.e., a signal that does not correspond to the actual action taken by the player. While r is the probability of an error for only two players. It represents the possibility that two players receive incorrect signals simultaneously. Then μ denotes the possibility of errors from three players. It likely refers to the overall error rate in the signaling mechanism, considering errors from all players.

By summing these individual and simultaneous error probabilities, $3(\varepsilon + r) + \mu$, the total probability of any error occurring is obtained. Therefore, to find the probability that no player makes a mistake this total error probability is subtracted from 1. The probability that no player will make a mistake is $\tau = 1 - (3(\varepsilon + r) + \mu$.

In three players we have 8 states which are (C, CC), (C, CD), (C, DC), (D, CC), (D, CD), (D, DC), (D, DD) respectively. For example, the state (C, CD) means that players 1 and 2 cooperate while player 3 defects.

For example,

$$\begin{aligned}
\phi\,(ggg \mid CCC) &= 1 - (3(\varepsilon + r) + \mu), \\
\phi\,(ggb \mid CCC) &= \phi\,(gbg \mid CCC) = \phi\,(bgg \mid CCC) = \varepsilon, \\
\phi\,(gbb \mid CCC) &= \phi\,(bbg \mid CCC) = \phi\,(bgb \mid CCC) = r, \\
\phi\,(bbb \mid CCC) &= \mu.
\end{aligned}$$

In games where players have a long-term relationship, they have more incentive to cooperate and build trust. This is because they know that they will be interacting with each other in the future, and they want to maintain a good relationship. In this case, the errors in the signaling mechanism can be more damaging because they can lead to players distrusting each other and making mistakes about each other's actions.

In each round, the expected payoff for player i (denoted as $f_i(a)$) is calculated based on his own action a_i and the distribution of signals x_i. A player's realized reward, or $G_i(a_i, x_i)$, is determined by both his actions and the signals he gets during a given round, as was previously indicated.

$$\begin{aligned}
G_i\,(C \mid gg) &= R, \quad G_i\,(C \mid gb) = G_i\,(C \mid bg) = K, \\
G_i\,(C \mid bb) &= S, \\
G_i\,(D \mid gg) &= T, \quad G_i\,(D \mid gb) = G_i\,(D \mid bg) = L, \\
G_i\,(D \mid bb) &= P.
\end{aligned}$$

The expected payoff $f_i(a)$ can be calculated as follows

$$f_i(a) = \sum_{x_i} [G_i(a_i, x_i)\phi(x|a)]. \tag{3.1}$$

The players' profit can depend on the signals they receive. In contrast to effect observation, where all players directly observe the actions of others, private observation involves receiving signals that may not fully reveal the opponents' actions.

The expected result is only determined by the activity profile a, regardless of the signal profile x_i. Hence, the expected return matrix is

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} CC & CD & DC & DD \end{array} \\ \begin{pmatrix} R_\xi & K_\xi & K_\xi & S_\xi \\ T_\xi & L_\xi & L_\xi & P_\xi \end{pmatrix} \end{array}. \tag{3.2}$$

According to Eq.(3.1) $R_\xi, S_\xi, T_\xi, L_\xi.K_\xi$ and P_ξ are derived as

$$\begin{aligned}
P_\xi &= G(D \mid gg)\,\phi(bbb \mid DDD) + G(D \mid gb)\,\phi(gbg \mid DDD) \\
&\quad + G(D \mid gb)\,\phi(ggb \mid DDD) \\
&\quad + G(D \mid bb)\,\phi(ggb \mid DDD) + G(D \mid gg)\,\phi(gbb \mid DDD) \\
&\quad + G(D \mid gb)\,[\phi(gbb \mid DDD) + \phi(bgb \mid DDD)] \\
&\quad + G(D \mid bb)\,\phi(ggg \mid DDD) \\
&= T\mu + 2\varepsilon L + P\varepsilon + Tr + 2rL + P\tau \\
&= T(r+\mu) + 2L(r+\varepsilon) + P(\tau+\varepsilon) \\
&= T(r+\mu) + 2L(r+\varepsilon) + P(1-\mu-3r-2\varepsilon),
\end{aligned}$$

$$\begin{cases}
R_\xi = 2k(r+\varepsilon) + S(r+\mu) + R(1-\mu-3r-2\varepsilon), \\
K_\xi = (r+\varepsilon)(R-2K+S) + K, \\
S_\xi = 2k(r+\varepsilon) + R(r+\mu) + S(1-\mu-3r-2\varepsilon), \\
T_\xi = 2L(r+\varepsilon) + P(r+\mu) + T(1-\mu-3r-2\varepsilon), \\
L_\xi = (r+\varepsilon)(T+P-2L) + L, \\
P_\xi = (r+\mu)T + 2L(r+\varepsilon) + P(1-\mu-3r-2\varepsilon).
\end{cases} \tag{3.3}$$

A player's value for future benefits relative to immediate rewards is gauged by the discount factor. A greater discount factor implies that a player places more importance on future wins, whereas a lower discount factor suggests that a player values recent wins more.

The chance of collaboration is influenced by the deduction factor in the IPD game with special observation. As participants are more ready to forego some immediate advantages in exchange for greater rewards in the future, cooperation is more likely when the discount factor is higher.

3.2. **Determinant form of expected payoff in the RPD game.** Consider player i that adopts memory-one strategies, with which they can use only the outcomes of the last round to decide the action to be submitted in the current round. A memory-one strategy is specified by a 7−tuple; X 's strategy is given by a combination of

$$p = (J_1, J_2, J_3, J_4, J_5, J_6, J_7; J_0), \tag{3.4}$$

where $0 \leq J_{ij} \leq 1$, $j \in \{0, 1, 2, 3, \ldots, 8\}$. The subscripts $1, 2, 3,$and 8 of p mean previous outcomes $C_{gg}, C_{gb}, C_{bg}, C_{bb}, D_{gg}, D_{gb}, D_{bg}$, and D_{bb}, respectively. In Eq. (3.4)), J_1 is the conditional probability that X cooperates when X cooperated and observed signal gg in the last round, J_2 is the conditional probability that X cooperates when X cooperated and observed signal gb in the last round, J_3 is the conditional probability that X cooperates and when

X cooperated and observed signal bg in the last round. Similarly to J_i for $i = 4, 5, ..., 8$ for denotes the corresponding conditional probabilities. Finally, p_0 is the probability that X cooperates in the first round. Similarly, z,Y' strategies.

In each round of the game, the reward subtracted from player i for a given action profile $a(t)$ at time t, where t ranges from 0 to infinity, is calculated as δ^t from the original payoff $f_i(a(t))$. In Eq.(3.1), the discounting process recognizes that players may value immediate rewards more than future rewards. For player i, the average discounted payoff is denoted by

$$n_\alpha = (1-\delta)\sum_{t=0}^{\infty} \delta^t f_i(a(t)). \tag{3.5}$$

Define $v(t) = (v_1(t), v_2(t), v_3(t), \ldots, v_8(t))$ as the stochastic state of three players in round t where the subscripts $1, 2, 3, \ldots,$ and 8 of v imply the stochastic states (C, CC), (C, CD), (C, DC), (D, CC), (D, CD), (D, DC), (D, DD) respectively.

The expected payoff is a weighted average of the possible payoffs, which are determined by the probability of occurrence of each situation given in Eq.(2.6). Let $V^t = (1-\delta)V(0)\,(I - \delta A)^{-1}$ be the mean distribution of $V(t)$. In addition, we define

$$\mathrm{M}_0 = \begin{bmatrix}
\jmath_0\kappa_0 z_0 & \jmath_0\kappa_0(1-z_0) & \jmath_0 z_0(1-\kappa_0) & \jmath_0(1-\kappa_0)(1-z_0) & \kappa_0 z_0(1-\jmath_0) & \kappa_0(1-\jmath_0)(1-z_0) & z_0(1-\jmath_0)(1-\kappa_0) & (1-\jmath_0)(1-\kappa_0)(1-z_0)\\
\jmath_0\kappa_0 z_0 & \jmath_0\kappa_0(1-z_0) & \jmath_0 z_0(1-\kappa_0) & \jmath_0(1-\kappa_0)(1-z_0) & \kappa_0 z_0(1-\jmath_0) & \kappa_0(1-\jmath_0)(1-z_0) & z_0(1-\jmath_0)(1-\kappa_0) & (1-\jmath_0)(1-\kappa_0)(1-z_0)\\
\jmath_0\kappa_0 z_0 & \jmath_0\kappa_0(1-z_0) & \jmath_0 z_0(1-\kappa_0) & \jmath_0(1-\kappa_0)(1-z_0) & \kappa_0 z_0(1-\jmath_0) & \kappa_0(1-\jmath_0)(1-z_0) & z_0(1-\jmath_0)(1-\kappa_0) & (1-\jmath_0)(1-\kappa_0)(1-z_0)\\
\jmath_0\kappa_0 z_0 & \jmath_0\kappa_0(1-z_0) & \jmath_0 z_0(1-\kappa_0) & \jmath_0(1-\kappa_0)(1-z_0) & \kappa_0 z_0(1-\jmath_0) & \kappa_0(1-\jmath_0)(1-z_0) & z_0(1-\jmath_0)(1-\kappa_0) & (1-\jmath_0)(1-\kappa_0)(1-z_0)\\
\jmath_0\kappa_0 z_0 & \jmath_0\kappa_0(1-z_0) & \jmath_0 z_0(1-\kappa_0) & \jmath_0(1-\kappa_0)(1-z_0) & \kappa_0 z_0(1-\jmath_0) & \kappa_0(1-\jmath_0)(1-z_0) & z_0(1-\jmath_0)(1-\kappa_0) & (1-\jmath_0)(1-\kappa_0)(1-z_0)\\
\jmath_0\kappa_0 z_0 & \jmath_0\kappa_0(1-z_0) & \jmath_0 z_0(1-\kappa_0) & \jmath_0(1-\kappa_0)(1-z_0) & \kappa_0 z_0(1-\jmath_0) & \kappa_0(1-\jmath_0)(1-z_0) & z_0(1-\jmath_0)(1-\kappa_0) & (1-\jmath_0)(1-\kappa_0)(1-z_0)\\
\jmath_0\kappa_0 z_0 & \jmath_0\kappa_0(1-z_0) & \jmath_0 z_0(1-\kappa_0) & \jmath_0(1-\kappa_0)(1-z_0) & \kappa_0 z_0(1-\jmath_0) & \kappa_0(1-\jmath_0)(1-z_0) & z_0(1-\jmath_0)(1-\kappa_0) & (1-\jmath_0)(1-\kappa_0)(1-z_0)\\
\jmath_0\kappa_0 z_0 & \jmath_0\kappa_0(1-z_0) & \jmath_0 z_0(1-\kappa_0) & \jmath_0(1-\kappa_0)(1-z_0) & \kappa_0 z_0(1-\jmath_0) & \kappa_0(1-\jmath_0)(1-z_0) & z_0(1-\jmath_0)(1-\kappa_0) & (1-\jmath_0)(1-\kappa_0)(1-z_0)
\end{bmatrix}. \tag{3.6}$$

The transition matrix with errors is given by W

$$W_{i(jk)}^{xyz} = \tau xyu + \varepsilon \tilde{A}_x^{yz} B_x^{yz} + r[x\tilde{y}\tilde{u} + \tilde{x}y\tilde{u} + \tilde{x}\tilde{y}u] + \mu\tilde{x}\tilde{y}\tilde{u}. \tag{3.7}$$

The variable x signifies the probability of player α being in state $i(jk)$. In matrix A, (jk) represents the column, while i represents the row. Similarly, the variables y and u represent the probabilities for player β and γ respectively. Where

$$\tilde{A}_x^{yz} = \begin{bmatrix} \tilde{x} & \tilde{y} & \tilde{u} \end{bmatrix}, B_x^{yz} = \begin{bmatrix} yu \\ xu \\ xy \end{bmatrix}, \tag{3.8}$$

$$x\,(yu) = a_{i(jk)}, \tilde{x}\,(\tilde{y}\tilde{u}) = \tilde{a}_{i(jk)}, \qquad i,j,k = 1,2,\dots,8. \tag{3.9}$$

Taking

$$x = \jmath_i \text{ or } 1-\jmath_i,$$
$$y = \kappa_j \text{ or } 1-\kappa_j,$$
$$u = z_k \text{ or } 1-z_k.$$

Let ω be defined as the set $\{i,j,k\}$, representing the counters for $\jmath, \kappa$, and z respectively. We partition the set $\{1,2,\dots,8\}$ into two subsets: $\mu_1 = \{1,2,3,4\}$ and $\mu_2 = \{5,6,7,8\}$. Let F be the set $\{x,y,u\}$, denoting the probabilities for the player. Additionally, let $\tilde{F}$ be the set $\{\tilde{x},\tilde{y},\tilde{u}\}$ for their respective counterparts. The set ω of counters for $\jmath$, κ and z, respectively.

$$\begin{aligned} \tilde{F} &= \sum_{l=1}^{2} \delta_l \sum_{\mu_l} F_{\mu_l}, \qquad \mu_l \neq \omega, \\ \delta_l &= \begin{cases} 1, & i \in \mu_l \;\wedge\; i \neq l, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{3.10}$$

According to Eq. (2.8),

$$V^T(1-\delta A) = (1-\delta)V^T M_0. \tag{3.11}$$

Equation (3.11) and

$$A^T \equiv \delta A + (1-\delta)M_0 - I \tag{3.12}$$

yield

$$V^T A^T = 0. \tag{3.13}$$

We combine identical results when we minimize the number of possible outcomes in a game. This indicates that the participants don't care which of these events happens. The game will be made simpler and easier to analyses by lowering the number of possible outcomes. Now, let's examine the dot product of the steady-state distribution vector V with an arbitrary six-dimensional vector $f = (f_1, f_2, f_3, f_4, f_5, f_6)$, represented as: Due to assuming a symmetric game matrix, we can write XCD as XDC, where X could be either C or D.

$$V \cdot f = D(\jmath, \kappa, z, f) = \begin{vmatrix} A_1 & A_2 & A_3 & A_4 & A_5 & A_6 \end{vmatrix},$$

where

$$\mathcal{A}_1 = \begin{array}{c} W^{xyz}_{1(11)} + \jmath_0\kappa_0 z_0(1-\delta) - 1 \\ W^{xyz}_{2(11)} + \jmath_0\kappa_0(1-z_0)(1-\delta) \\ W^{xyz}_{3(11)} + \jmath_0(1-\kappa_0)(1-z_0)(1-\delta) \\ W^{xyz}_{4(11)} + (1-\jmath_0)\kappa_0 z_0(1-\delta) \\ W^{xyz}_{5(11)} + (1-\jmath_0)\kappa_0(1-z_0)(1-\delta) \\ W^{xyz}_{6(11)} + (1-\jmath_0)(1-\kappa_0)(1-z_0)(1-\delta) \end{array}$$

$$\mathcal{A}_2 = \begin{array}{c} W^{xyz}_{1(12)} + \jmath_0\kappa_0 z_0(1-\delta) \\ W^{xyz}_{2(12)} + \jmath_0\kappa_0(1-z_0)(1-\delta) \\ W^{xyz}_{3(12)} + \jmath_0(1-\kappa_0)(1-z_0)(1-\delta) \\ W^{xyz}_{4(12)} + (1-\jmath_0)\kappa_0 z_0(1-\delta) \\ W^{xyz}_{5(12)} + (1-\jmath_0)\kappa_0(1-z_0)(1-\delta) \\ W^{xyz}_{6(12)} + (1-\jmath_0)(1-\kappa_0)(1-z_0)(1-\delta) \end{array}$$

$$\mathcal{A}_3 = \begin{array}{c} \sigma(\varsigma\jmath_1 + \eta(2\jmath_2 + \jmath_3)) + \jmath_0(1-\sigma) - 1 \\ \sigma(\varsigma\jmath_2 + \eta(\jmath_1 + \jmath_2 + \jmath_3)) + \jmath_0(1-\sigma) - 1 \\ \sigma(\varsigma\jmath_3 + \eta(\jmath_1 + 2\jmath_2)) + \jmath_0(1-\sigma) \\ \sigma(\varsigma\jmath_4 + \eta(2\jmath_5 + \jmath_6)) + \jmath_0(1-\sigma) \\ \sigma(\varsigma\jmath_5 + \eta(\jmath_4 + \jmath_5 + \jmath_6)) + \jmath_0(1-\sigma) - 1 \\ \sigma(\varsigma\jmath_6 + \eta(\jmath_4 + 2\jmath_5)) + \jmath_0(1-\sigma) \end{array}$$

$$\mathcal{A}_4 = \begin{array}{c} \sigma(\varsigma\kappa_1 + \eta(2\kappa_2 + \kappa_3)) + \kappa_0(1-\sigma) - 1 \\ \sigma(\varsigma\kappa_2 + \eta(\kappa_1 + \kappa_2 + \kappa_3)) + \kappa_0(1-\sigma) - 1 \\ \sigma(\varsigma\kappa_3 + \eta(\kappa_1 + 2\kappa_2)) + \kappa_0(1-\sigma) \\ \sigma(\varsigma\kappa_4 + \eta(2\kappa_5 + \kappa_6)) + \kappa_0(1-\sigma) - 1 \\ \sigma(\varsigma\kappa_5 + \eta(\kappa_4 + \kappa_5 + \kappa_6)) + \kappa_0(1-\sigma) \\ \sigma(\varsigma\kappa_6 + \eta(\kappa_4 + 2\kappa_5)) + \kappa_0(1-\sigma) \end{array}$$

$$\mathcal{A}_5 = \begin{array}{c} \sigma(\varsigma z_1 + \eta(2z_2 + z_3)) + z_0(1-\sigma) - 1 \\ \sigma(\varsigma z_2 + \eta(z_1 + z_2 + z_3)) + z_0(1-\sigma) - 1 \\ \sigma(\varsigma z_3 + \eta(z_1 + 2z_2)) + z_0(1-\sigma) \\ \sigma(\varsigma z_4 + \eta(2z_5 + z_6)) + z_0(1-\sigma) - 1 \\ \sigma(\varsigma z_5 + \eta(z_4 + z_5 + z_6)) + z_0(1-\sigma) \\ \sigma(\varsigma z_6 + \eta(z_4 + 2z_5)) + z_0(1-\sigma) \end{array}$$

$$\mathcal{A}_6 = \begin{matrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{matrix}$$

An arbitrary vector's dot product with its mean distribution's V. The expected payoff

$$\begin{aligned} n_\alpha &= V.N_\alpha = \frac{D(\jmath,\kappa,z,n_\alpha)}{D(\jmath,\kappa,z,1)}, \\ n_\beta &= V.N_\beta = \frac{D(\jmath,\kappa,z,n_\beta)}{D(\jmath,\kappa,z,1)}, \\ n_\gamma &= V.N_\gamma = \frac{D(\jmath,\kappa,z,n_\gamma)}{D(\jmath,\kappa,z,1)}. \end{aligned} \tag{3.14}$$

The analysis involves calculating players' per-round expected payoffs for $0 < \delta \leq 1$ using determinant calculations. $\delta = 1$ represents the scenario of undiscounted future payoffs. In the context of the IPD game, it has been found that despite having errors in observation and considering the discount factor, only those strategies exist that establish a linear correlation among the three players' payoffs.

$$\begin{cases} \sigma(\varsigma \jmath_1 + \eta(2\jmath_2 + \jmath_3)) + \jmath_0(1-\sigma) - 1 = \alpha R_\xi + \beta R_\xi + \delta R_\xi + \gamma, \\ \sigma(\varsigma \jmath_2 + \eta(\jmath_1 + \jmath_2 + \jmath_3)) + \jmath_0(1-\sigma) - 1 = \alpha K_\xi + \beta K_\xi + \delta T_\xi + \gamma, \\ \sigma(\varsigma \jmath_3 + \eta(\jmath_1 + 2\jmath_2) + \jmath_0(1-\sigma) - 1 = \alpha S_\xi + \beta L_\xi + \delta L_\xi + \gamma, \\ \sigma(\varsigma \jmath_4 + \eta(2\jmath_5 + \jmath_6)) + \jmath_0(1-\sigma) = \alpha T_\xi + \beta K_\xi + \delta K_\xi + \gamma, \\ \sigma(\varsigma \jmath_5 + \eta(\jmath_4 + \jmath_5 + \jmath_6)) + \jmath_0(1-\sigma) = \alpha L_\xi + \beta L_\xi + \delta S_\xi + \gamma, \\ \sigma(\varsigma \jmath_6 + \eta(\jmath_4 + 2\jmath_5)) + \jmath_0(1-\sigma) = \alpha P_\xi + \beta P_\xi + \delta P_\xi + \gamma. \end{cases} \tag{3.15}$$

4. When there are errors, extortion no longer exists

ZD strategies are well-known subsets with important traits, and extortion strategies fit into this category. Always, extortion doesn't lose to any opponent. An opponent trying to improve his reward forces him to cooperate [27]. In the context of large populations that evolve over time, it is emphasized that groups that adopt strategies of cooperation are more successful compared to those groups that pursue strategies of extortion. What is meant is that in large, complex environments that evolve over time, groups that cooperate and work together appear to be more successful in the end. This is because cooperation can lead to a win-win solution that contributes to the overall susceptibility and success of the group, while extortion and selfish behavior can lead to deteriorating relationships and group disintegration over time.

Substituting $\alpha = \phi$, $\chi = -\frac{\beta}{\alpha}$, and $\gamma = \phi(2\chi - 1)\varpi$ into equation (3.15), we obtain

$$\begin{cases} \sigma\big(\varsigma\jmath_1 + \eta(2\jmath_2 + \jmath_3)\big) + \jmath_0(1-\sigma) - 1 = \phi(1-2\chi)(R_\xi - \varpi), \\ \sigma\big(\varsigma\jmath_2 + \eta(\jmath_1 + \jmath_2 + \jmath_3)\big) + \jmath_0(1-\sigma) - 1 = \phi\big[(K_\xi - \varpi) - \chi(K_\xi + T_\xi - 2\varpi)\big], \\ \sigma\big(\varsigma\jmath_3 + \eta(\jmath_1 + 2\jmath_2)\big) + \jmath_0(1-\sigma) - 1 = \phi\big[S_\xi - \varpi - 2\chi(L_\xi - \varpi)\big], \\ \sigma\big(\varsigma\jmath_4 + \eta(2\jmath_5 + \jmath_6)\big) + \jmath_0(1-\sigma) = \phi\big[T_\xi - \varpi - 2\chi(K_\xi - \varpi)\big], \\ \sigma\big(\varsigma\jmath_5 + \eta(\jmath_4 + \jmath_5 + \jmath_6)\big) + \jmath_0(1-\sigma)\phi\big[L_\xi - \varpi - \chi(L_\xi + S_\xi - 2\varpi)\big], \\ \sigma\big(\varsigma\jmath_6 + \eta(\jmath_4 + 2\jmath_5)\big) + \jmath_0(1-\sigma) = \phi(1-2\chi)(P_\xi - \varpi), \end{cases} \tag{4.1}$$

where $\varpi = P_\xi$ with $1 \leq \chi < \infty$ represents extortion, $\varsigma = \tau + 9r + 6\varepsilon$ and $\eta = 9\mu + 6r + \varepsilon$. Eq.(4.1) is obtained by solving the stationary distribution of the Markov chain and substituting it into the long-run payoff expression. Concretely, we proceed as follows. From the transition matrix M we solve the stationary equation $\pi M = \pi$ together with the normalization $\sum_i \pi_i = 1$. The parameter χ represents the degree of correlation between the outcomes of the players in the game. This parameter reflects how similar or related players' rewards are based on the choices they make. The parameter ϖ indicates the outcome that the ZD strategy will achieve against itself. Specifically, if the player decides to follow the ZD strategy against himself, he will have the ability to get a specific reward representing the parameter ϖ. This indicates how the ZD strategy interacts with itself and how player rewards can be affected based on these choices. In the next round, Player α will employ strategy $\tilde{p}$ and adapt his choice to correspond with the last column of the previous matrix in A.

$$\begin{cases} \sigma(\varsigma\jmath_1 + \eta(2\jmath_2 + \jmath_3)) + \jmath_0(1-\sigma) - 1 = \alpha R_\xi + \gamma, \\ \sigma(\varsigma\jmath_2 + \eta(\jmath_1 + \jmath_2 + \jmath_3)) + \jmath_0(1-\sigma) - 1 = \alpha k_\xi + \gamma, \\ \sigma(\varsigma\jmath_3 + \eta(\jmath_1 + 2\jmath_2) + \jmath_0(1-\sigma) - 1 = \alpha S_\xi + \gamma, \\ \sigma(\varsigma\jmath_4 + \eta(2\jmath_5 + \jmath_6)) + \jmath_0(1-\sigma) = \alpha T_\xi + \gamma, \\ \sigma(\varsigma\jmath_5 + \eta(\jmath_4 + \jmath_5 + \jmath_6)) + \jmath_0(1-\sigma) = \alpha L_\xi + \gamma, \\ \sigma(\varsigma\jmath_6 + \eta(\jmath_4 + 2\jmath_5)) + \jmath_0(1-\sigma) = \alpha P_\xi + \gamma, \end{cases} \tag{4.2}$$

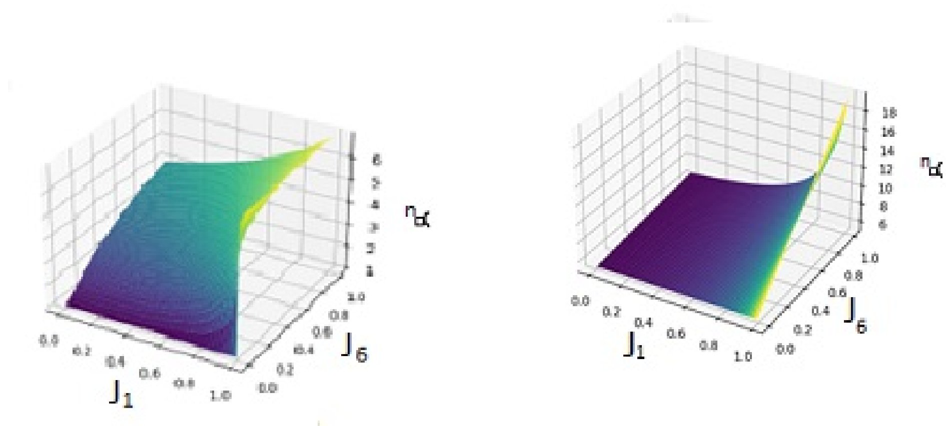
when we solve Eq.(4.2), we obtain

$$\begin{cases} \jmath_2 = & \jmath_1 + \dfrac{\alpha(R_\xi - k_\xi)}{\sigma(\eta-\varsigma)}, \\ \jmath_3 = & \jmath_1 + \dfrac{\alpha(R_\xi - S_\xi)}{\sigma(\eta-\varsigma)}, \\ \jmath_4 = & \jmath_6 - \dfrac{\alpha(T_\xi - P_\xi)}{\sigma(\eta-\varsigma)}, \\ \jmath_5 = & \jmath_6 - \dfrac{\alpha(L_\xi - P_\xi)}{\sigma(\eta-\varsigma)}, \end{cases} \tag{4.3}$$

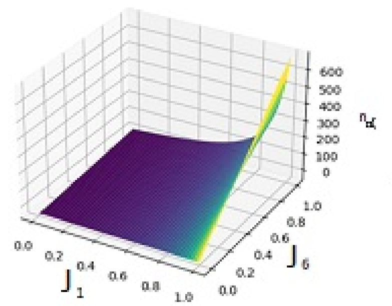
$$\alpha = \frac{[1-(\jmath_1-\jmath_6)(\varsigma+3\eta)]\sigma(\eta-\varsigma)}{(R_\xi-P_\xi)(3\eta-\sigma(\eta-\varsigma))+\eta(T_\xi-S_\xi-2k_\xi-2L_\xi)}, \tag{4.4}$$

$$n_\alpha = \frac{[\jmath_1(\varsigma+3\eta)-1+\jmath_0(1-\sigma)][(R_\xi-P_\xi)(3\eta-\sigma(\eta-\varsigma))+\eta(T_\xi-S_\xi-2k_\xi-2L_\xi)]}{\sigma(\eta-\varsigma)[(\varsigma+3\eta)(\jmath_1-\jmath_6)-1]} - \frac{\eta(3R_\xi-S_\xi-2k_\xi)}{\sigma(\eta-\varsigma)} + R. \tag{4.5}$$

The condition $0 \leq \jmath_i \leq 1$, $i = \{1,2,\ldots,6\}$ is satisfied in equalizer in Eq.(4.3). In this case, σ, ς, and η are likely key parameters that affect whether equalizer strategies can be employed or are effective in each context. The conditions under which equalizer strategies are feasible or optimal could be influenced by the values of these parameters.



(A) No noise (B) Low noise



(c) High noise

Figure 1. Payoff of all strategies in three players

Figure 1: Illustration of pinning strategies in a three-player game under varying levels of noise.It shows that, as the noise level increases, the payoff advantage of extortion strategies disappears. The shaded region in each subfigure represents the feasible payoff outcomes for player β when player α employs a pinning strategy. The accuracy and effectiveness of such strategies diminish as noise increases. Noise in the system can arise from three sources: (i) observation error ε, where a player's action is misperceived; (ii) perception error γ, where a received signal is misinterpreted; and (iii) implementation error μ, where a player intends one action but executes another.

(a) No noise: $\varepsilon = 0, r = 0, \mu = 0$.
(b) Low noise: $\varepsilon = 0.005, r = 0.055, \mu = 0.07$.
(c) High noise: $\varepsilon = 0.1, r = 0.1, \mu = 0.3$.

These numerical experiments confirm that the analytical ZD conditions remain valid even under varying levels of noise. The simulations clearly show that increasing observation errors weakens the enforcement of linear payoff relations and eliminates extortion behavior. Overall, the numerical patterns align with theoretical predictions and demonstrate the robustness and limitations of ZD strategies in noisy three-player settings.

5. ZD in a strategic form of TFT, ALLC and ALLD with observation

While no one strategy fits all scenarios perfectly, a remarkably effective approach across diverse situations is the straightforward Tit For Tat (TFT) strategy. For the TFT strategy, the player cooperates initially and subsequently mirrors the opponent's actions. Since there are many strategies in the game (I3PD), we concentrate on the six strategies known as TFT-1, TFT-2, TFT-3, TFT-4, ALLC, and ALLD, then calculate the payoff values by observing the error and displaying it in tabular form. A thorough examination was carried out to analyze the choice of tactics and rivalry between a range of generalized ZD tactics and the conventional tactics of constant cooperation (AllC), constant defection (ALL D), (TFT), and win-stay-lose-shift (WSLS) strategies within an IPD [20].

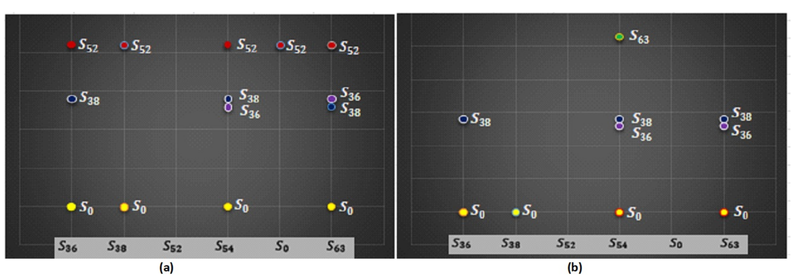
The TFT strategy stands out for its balance as it first cooperates and tracks the opponent's response, making it gentle and vindictive. It is also difficult to exploit because any separation from cooperation will find itself returning. When used, the player begins to cooperate in the first round, and if it recognizes that the opponent has separated, it will withdraw in the next round. After noticing the opponent's cooperation again, he prepares to cooperate again. Studies show that the Tit For Tat strategy contributes to enhancing cooperation in non-traditional groups, and it has been tried by Axelrod as a savior for cooperation in the evolution of strategies [26].

We concentrate on the six strategies known as TFT-1, TFT-2, TFT-3, TFT-4, ALLC, and ALLD which are S_{36}, S_{38}, S_{52}, S_{54}, S_{63} and S_0 respectively, then calculate the payoff values by observing the error and displaying it in tabular form.

Definition 5.1. (Dominating Strategy) A strategy S_i is said to be dominating for player i if, for every possible strategy profile of the other players, S_i yields a payoff that is greater than or equal to the payoff obtained from any alternative strategy $\acute{S}_i$. Formally,

$$u_i\left(S_i, S_{-i}\right) \geq u_i\left(\acute{S}_i, S_{-i}\right), \quad \forall S_i,\ \acute{S}_i \neq S_i.$$

In this study, "dominating strategies" refer to those strategies that consistently maximize the player's payoff under the noisy observation model and across the examined parameter space.



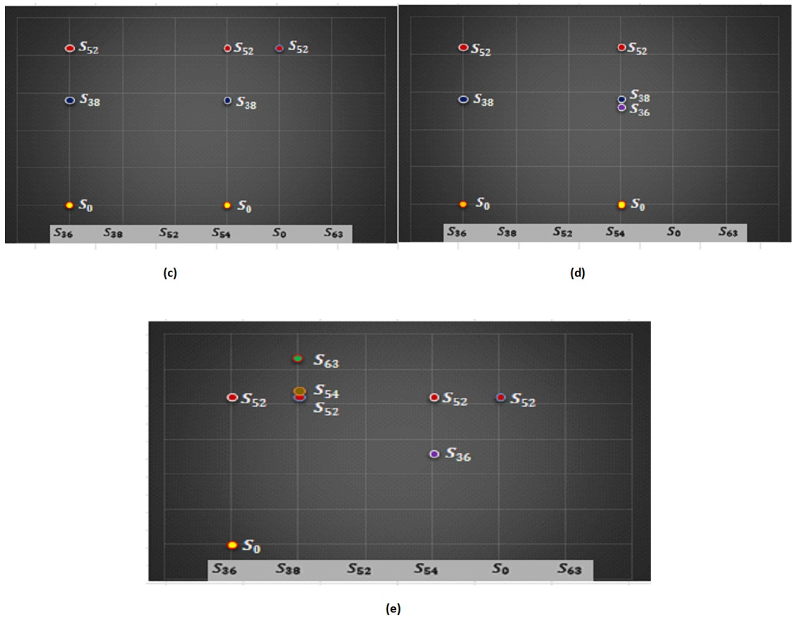
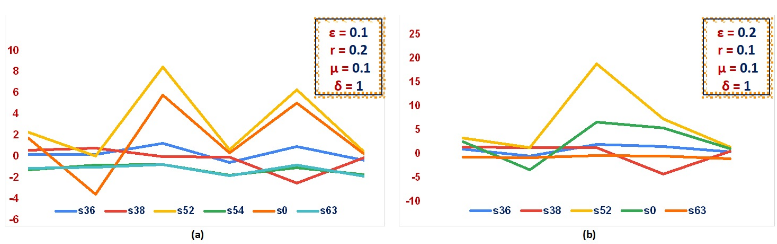
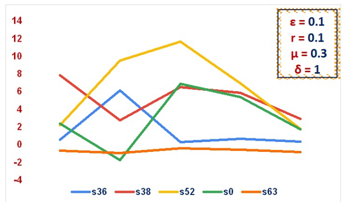


FIGURE 2. Dominating strategies for general conditions

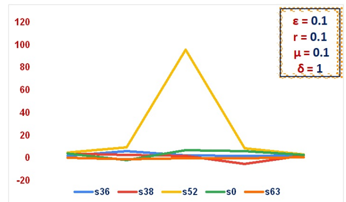
Figure 2 depicts the behavior of dominating strategies. It shows that no strategy could defeat strategy S_{52}, and it could also defeat four strategies, including the All-D S_0. As a result, strategy S_{52} compelled us to conclude that it performed admirably in this study. There are other techniques with better performance, such as S_0 and S_{38}.

Some feeble strategies make it possible for alternative strategies to infiltrate their defenses and lack the capacity to overpower them. Surprisingly, Strategy S_0 outperforms five other strategies, with only two other strategies outperforming S_0. Poor strategies S_{54} and S_{63} only defeated two strategies at most and were defeated by five others. Only two strategies, S_{52} and S_{36}, invaded the two powerful strategies S_0, and S_{38}.

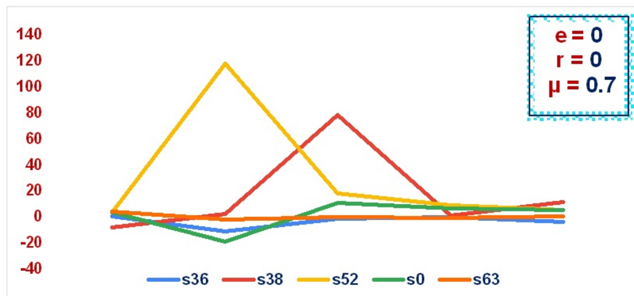




(c)



(d)



(e)

FIGURE 3. Dominating strategies for general conditions

Figure 3 illustrates the outcome or result of each strategy in comparison to itself and all other strategies. In each of the five sub-figures, the S_{52} strategy is accompanied by five additional strategies ($S_{36}, S_{38}, S_{54}, S_0, S_{63}$). Each subfigure (a-e) depicts the effective strategy S_{52} with different error values to demonstrate how it outperforms other strategies.

When S_{38} is fixed, S_{52} is the best against S_{54}, S_0, S_{63} but S_{36} and S_{38} are the best against S_{52}. And in (b) and (c), when S_{38} is fixed, S_{52} is the best against $S_{36}, S_{54}, S_0, S_{63}$ but S_{38} is the best against S_{52}. Also in (d), when S_{38} is fixed, S_{52} is the best against $S_{36}, S_{38}, S_{54}, S_0, S_{63}$. However, in (e), when S_{38} remains constant, S_{52} emerges as the optimal strategy against S_{36} and S_0. Conversely, S_{38} and S_{63} prove to be the most effective strategies against S_{52}. The strategies S_{52} and S_0 have the same payoff with S_{36}. The S_{52} strategy has established itself as a formidable strategy. Due to the fact that strategy S_{52} does not engage in cooperation unless it deceives its adversary, it can be considered as a flawed strategy. In the context of these competitions, strategy S_{63} exhibited the poorest performance.

6. Conclusion

In conclusion, this study has illuminated the significance of ZD strategies within a three-player repeated game framework, showcasing their potential to manipulate reward dynamics even in the presence of observation errors. The findings emphasize that ZD strategies can establish linear relationships between players' payoff, a phenomenon that holds true despite the complex interplay of strategies and imperfect observations.

The analysis underscores the impact of discounting factors and observational inaccuracies on the effectiveness of ZD strategies. By considering these critical factors, the study has deepened our understanding of the strategic dynamics in the IPD game. The numerical example from results presented in a 3x3 PD scenario provides empirical evidence of the applicability and outcomes of ZD strategies under realistic conditions.

The insights gleaned from this research contribute to a more comprehensive understanding of strategic decision-making and cooperation patterns in repeated game scenarios. Furthermore, it underscores the importance of considering the intricacies of observation errors and discounting factors when assessing the feasibility and influence of ZD strategies. As the realm of game theory continues to evolve, the study of ZD strategies and their implications remains a promising avenue for further research and exploration in the complex landscape of strategic interactions.

Finally, upon comparing the four variations of TFT, we can assert that TFT 3 demonstrates a higher level of adaptability compared to TFT 4. This is attributed to the fact that in TFT 3, the player can transition from state D to C after one of their opponents played C in the previous round, whereas TFT 4 requires both opponents to play C simultaneously. Similarly, TFT 2 exhibits more flexibility than TFT 1 due to the same reasoning. Moreover, TFT 4's strategy proves to be more lenient than TFT 2, as TFT 4 can skip state C and proceed to D only if both opponents have played D in the previous round. In contrast, TFT 2 will penalize both opponents if either of them played D. In addition, TFT strategy shows that 3 has a larger tolerance than TFT 1 for the same reason.

Regarding dominant strategies, S_{52} consistently outperforms four other strategies, including All-D S_0, though it is considered defective as it only cooperates when deceiving its opponent. Conversely, S_{63} consistently shows the weakest performance.

In conclusion, the study highlights the robustness and flexibility of ZD strategies in multiplayer games, even under observation errors and discounting effects. It also underscores the performance nuances of different strategic variants, identifying conditions in which specific strategies are either more effective or defective, thereby enhancing understanding of cooperation dynamics and strategic decision-making in repeated game scenarios.

List of Abbreviation

Abbreviation	**Full Meaning**
ZD	Zero-Determinant
IPD	Iterated Prisoner's Dilemma
PD	Prisoner's Dilemma
C	Cooperation (Choice/Action)
D	Defect (Choice/Action)
g	Signal for Cooperative Behavior
b	Signal for Defection
ALLC	All Cooperate (Strategy)
ALLD	All Defect (Strategy)
TFT	Tit-for-Tat (Strategy)
3IPD	Three-player Iterated Prisoner's Dilemma
R	Reward (Payoff)
S	Sucker's Payoff (Payoff)
T	Temptation (Payoff)
P	Punishment (Payoff)
α	The Central Player (in the three-player game model)
β	Another Player (in the three-player game model)
γ	The Third Player (in the three-player game model)
ε	Probability of error for one player
r	Probability of error for only two players
μ	Possibility of errors from three players
χ	Degree of correlation between the outcomes of the players in the game
ϖ	The outcome that the ZD strategy will achieve against itself

Acknowledgments: The author would like to express their gratitude to professor Dr. Essam El-Seidy, Mathematics Department, Faculty of Science, Al-Azhar University for suggesting the problem and critically reading the manuscript.

References

[1] S.M. Abd El-Salam, E. El-Seidy and A.R. Abdel-Malek, *Evaluating zero-determinant strategies' effects on cooperation and conflict resolution in repeated games*, Chaos, Solitons and Fractals, **174** (2023), 113779.

[2] D. Abreu, B. Brooks and Y. Sannikov, *Algorithms for stochastic games with perfect monitoring*, Econometrica, **88**(4) (2020), 1661–1695.

[3] D. Abreu and Y. Sannikov, *An algorithm for two-player repeated games with perfect monitoring*, Theoretical Economics, **9**(2) (2014), 313–338.

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

[5] E. Akin, *What you gotta know to play good in the iterated prisoner's dilemma*, Games, **6**(3) (2015), 175–190.

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

[7] M. Aoyagi, V. Bhaskar and G.R. Fréchette, *The impact of monitoring in infinitely repeated games: Perfect, public, and private*, American Economic Journal: Microeconomics, **11**(1) (2019), 1–43.

[8] M. Barlo, G. Carmona and H. Sabourian, *Repeated games with one-memory*, Journal of Economic Theory, **144**(1) (2009), 312–336.

[9] L. Becks and M. Milinski, *Extortion strategies resist disciplining when higher competitiveness is rewarded with extra gain*, Nature Communications, **10**(1) (2019), 783.

[10] A. Engel and A. Feigel, *Single equalizer strategy with no information transfer for conflict escalation*, Physical Review E, **98**(1) (2018), 012415.

[11] D. Fudenberg, D. Levine and E. Maskin, *The folk theorem with imperfect public information*, In: A Long-Run Collaboration on Long-Run Games, (2009), 231–273.

[12] D. Fudenberg, D.G. Rand and A. Dreber, *Slow to anger and fast to forgive: Cooperation in an uncertain world*, American Economic Review, **102**(2) (2012), 720–749.

[13] D. Hao, Z. Rong and T. Zhou, *xtortion under uncertainty: Zero-determinant strategies in noisy games*, Physical Review E, **91**(5) (2015), 052803.

[14] C. Hilbe, K. Chatterjee and M.A. Nowak, *artners and rivals in direct reciprocity*, Nature Human Behaviour, **2**(7) (2018), 469–477.

[15] C. Hilbe, A. Traulsen and K. Sigmund, *Partners or rivals? Strategies for the iterated prisoner's dilemma*, Games and Economic Behavior, **92** (2015), 41–52.

[16] C. Hilbe, B. Wu, A. Traulsen and M.A. Nowak, *Cooperation and control in multiplayer social dilemmas*, Proceedings of the National Academy of Sciences, **111**(46) (2014), 16425–16430.

[17] Q. Hu, S. Wang, P. Ma, X. Cheng, W. Lv and R. Bie, *Quality control in crowdsourcing using sequential zero-determinant strategies*, IEEE Transactions on Knowledge and Data Engineering, **32**(5) (2019), 998–1009.

[18] G. Ichinose and N. Masuda, *Zero-determinant strategies in finitely repeated games*, Journal of Theoretical Biology, **438** (2018), 61–77.

[19] M. Kandori, *Introduction to repeated games with private monitoring*, Journal of Economic Theory, **102**(1) (2002), 1–15.

[20] J. Liu, Y. Li, C. Xu and P.M. Hui, *Evolutionary behavior of generalized zero-determinant strategies in iterated prisoner's dilemma*, Physica A, **430** (2015), 81–92.

[21] G.J. Mailath and L. Samuelson, *Repeated games and reputations: Long-run relationships*, Oxford University Press, 2006.

[22] A. Mamiya and G. Ichinose, *Strategies that enforce linear payoff relationships under observation errors in repeated prisoner's dilemma game*, Journal of Theoretical Biology, **477** (2019), 63–76.

[23] A. Mamiya and G. Ichinose, *Zero-determinant strategies under observation errors in repeated games*, Physical Review E, **102**(3) (2020), 032115.

[24] Y. Mao, X. Xu, Z. Rong and Z.X. Wu, *The emergence of cooperation-extortion alliance on scale-free networks with normalized payoff*, Europhysics Letters, *122*(5) (2018), 50005.

[25] A. McAvoy and C. Hauert, *Autocratic strategies for alternating games*, Theoretical Population Biology, **113** (2017), 13–22.

[26] P.R. Milgrom, *Axelrod's (1984) The evolution of cooperation*, (1984), 305–309.

[27] W.H. Press and F.J. Dyson, *Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent*, Proceedings of the National Academy of Sciences, **109**(26) (2012), 10409–10413.

[28] T. Sekiguchi, Efficiency in repeated prisoner's dilemma with private monitoring, Journal of Economic Theory, **76**(2) (1997), 345–361.

[29] A.J. Stewart and J.B. Plotkin, *From extortion to generosity, evolution in the iterated prisoner's dilemma*, Proceedings of the National Academy of Sciences, **110**(38) (2013), 15348–15353.

[30] A. Szolnoki and M. Perc, *Evolution of extortion in structured populations*, Physical Review E, **89**(2) (2014), 022804.

[31] M.A. Taha and A. Ghoneim, *Zero-determinant strategies in repeated asymmetric games*, Applied Mathematics and Computation, **369** (2020), 124862.

[32] M. Tayer, *Possibility of cooperation: Studies in rationality and social change*, Printed in Great Britain at the University Press, Cambridge, 1987. ISBN 0 521 32793 8.

[33] J. Wang and J. Guo, *A synergy of punishment and extortion in cooperation dilemmas driven by the leader*, Chaos, Solitons and Fractals, **119** (2019), 263–268.

[34] Z. Wang, Y. Zhou, J.W. Lien, J. Zheng and B. Xu, *Extortion can outperform generosity in the iterated prisoner's dilemma*, Nature Communications, **7**(1) (2016), 11125.

[35] X. Xu, Z. Rong, Z. Tian and Z.X. Wu, *Timescale diversity facilitates the emergence of cooperation-extortion alliances in networked systems*, Neurocomputing, **350** (2019), 195–201.