How Does Misperception Affect Zero-Determinant Strategies In Iterated Prisoner’s Dilemma?

Zhaoyang Cheng  
Academy of Mathematics and Systems Science, Beijing

Guanpu Chen  
Academy of Mathematics and Systems Science, Beijing

Yiguang Hong (✉️ yghong@iss.ac.cn )  
Tongji University

Research Article

**Keywords:** strategies, utilities, misperception, cognition

**Posted Date:** December 22nd, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-1166862/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
How Does Misperception Affect Zero-Determinant Strategies in Iterated Prisoner’s Dilemma?

Zhaoyang Cheng1,2, Guanpu Chen1,2, and Yiguang Hong3,1,*

1Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Beijing, 100190, China
2School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing, 100190, China
3Department of Control Science and Engineering, Tongji University, Shanghai, 201804, China
*yghong@iss.ac.cn

ABSTRACT

Zero-determinant (ZD) strategies have attracted wide attention in Iterated Prisoner’s Dilemma (IPD) games, since the player equipped with ZD strategies can unilaterally enforce the two players’ expected utilities subjected to a linear relation. On the other hand, uncertainties, which may be caused by misperception, occur in IPD inevitably in practical circumstances. To better understand the situation, we consider the influence of misperception on ZD strategies in IPD, where the two players, player X and player Y, have different cognitions, but player X detects the misperception and it is believed to make ZD strategies by player Y. We provide a necessary and sufficient condition for the ZD strategies in IPD with misperception, where there is also a linear relationship between players’ utilities in player X’s cognition. Then we explore bounds of players’ expected utility deviation from a linear relationship in player X’s cognition with also improving its own utility.

Iterated Prisoner’s Dilemma (IPD) games have long been studied for understanding the evolution of cooperation and competition between players1–3. It is generated by a one-shot Prisoner’s Dilemma (PD) game between player X and player Y, where both of them choose to cooperate (c) or defect (d). Players’ utility matrix is shown in Table 1, where parameters \([T, R, P, S]\) of the PD game are constrained by \(T > R > P > S\) and \(2R > T + S\). Thus, mutual defection is the only Nash equilibrium, but mutual cooperation is the globally best outcome. In IPD games, the analysis of players’ utilities is complicated since players may promote cooperation through past actions. Fortunately, Press and Dyson7 proposed zero-determinant (ZD) strategies, a class of probabilistic memory-one strategies, where the player equipped with ZD strategies can unilaterally enforce the two players’ expected utilities subjected to a linear relation. Afterward, typical subsets of ZD strategies were widely studied in multiplayer games, incentive structures, and so on8–11. For example, the equalizer strategy7,12 is a special ZD strategy that can unilaterally set the opponent’s utility. Besides, the player who adopts extortion strategies7,9 can make that its utility is not lower than the opponent’s utility. Conversely, the generous strategy13 is another special ZD strategy that ensures that the utility of the player with generous strategies is not higher than the opponent’s utility, but it is dominant in the game.

Actually, uncertainty is always unavoidable in human interactions14, and there have been many models to describe uncertain circumstances in game theory, such as robust games, stochastic games, and hypergames15–17. Misperception is one of the most common uncertain phenomena. For example, in the Internet of Things, limited attention is a type of misperception, leading to bounded rationality and increasing cyber risks of the community18, and in cyber security problems, hackers may have a confused cognition of the system’s TCP/IP stack, which is known to the network administrator19. Moreover, players’ strategies may be influenced by uncertainty, which results in obvious deviation from opponents’ cognitions and attendant suspicion, such as the extenuating circumstances which consider intentions and outcomes in the legal system20,21, while players may misunderstand their opponents’ strategies, such as some companies relying on private monitoring instead of their opponents’ real actions22,23.

In fact, the condition for players to trust their cognition is crucial in games with misperception24,25. Particularly, misperception may spoil players’ cognition if others’ strategies are not consistent with their own anticipation, and moreover, it
may even ruin the balance or even lead to collapse of the model\textsuperscript{26}. For instance, in psychological experiments, participants’ doubts may affect the sponsor’s control\textsuperscript{27}. Actually, in IPD with misperception, players observe their strategies in sequence, which may deviate from their anticipated strategy sequences. They may doubt something if their opponents do not choose ZD strategies as they expect. Nevertheless, most existent works on ZD strategies in IPD with uncertainties, such as ZD strategies with observation errors\textsuperscript{22,21} or implementation errors\textsuperscript{20,28}, have paid less attention to strategies that maintain their opponents’ cognition.

Therefore, the motivation of this paper is to analyze how misperception affects players’ ZD strategies without causing their opponents’ suspicion. Specifically, we consider the case when player $X$ knows the misperception about the game, and player $Y$ believes that player $X$ prefers to make ZD strategies according to the original model without misperception. Then player $X$ tends to choose strategies consistent with its opponent’s anticipation, and meanwhile improve its own expected utility.

To this end, we find some conditions where player $X$ is able to achieve at least a linear relationship between players’ expected utilities without causing the opponent’s awareness of misperception. Additionally, misperception can bring a bounded deviation from the linear relationship between players’ expected utilities in player $X$’s cognition, which can be applied to player $X$’s strategy implementation. Further, player $X$ can utilize the misperception and take some benefits, such as improving the supremum or the infimum of its expected utility.

**Results**

**Models**

Consider an IPD game with misperception such as implementation errors and observation errors\textsuperscript{20,21,29}. Particularly, player $Y$’s cognition of the parameter is $\omega = [T_1, R_1, P_1, S_1]$ and player $X$’s cognition of the parameter is $\omega = [T_2, R_2, P_2, S_2]$, while player $X$ can notice the misperception. In each round, player $X$ choose a strategy from its strategy set $\Omega_X = \{p = [p_{cc}, p_{cd}, p_{dc}, p_{dd}]^T | p_{xy} \in [0, 1], xy \in \{cc, cd, dc, dd\}\}$, i.e., $p_{xy}$ is player $X$’s probability for cooperating with given previous outcome $xy \in \{cc, cd, dc, dd\}$. Similar to $\Omega_X$, player $Y$’s strategy set is $\Omega_Y = \{q = [q_{cc}, q_{cd}, q_{dc}, q_{dd}]^T | q_{xy} \in [0, 1], xy \in \{cc, cd, dc, dd\}\}$. According to Press and Dyson\textsuperscript{7}, this game can be characterized by a Markov chain with a state transition matrix $M = [M_{ij}]_{4 \times 4}$ (see “Notations” for details). Denote $v = [v_{cc}, v_{cd}, v_{dc}, v_{dd}]^T$ as the steady state probability vector of $M$ such that $v^T M = v^T$ and $\sum_{i=1}^4 v_i = 1$.

Let $S_X^0 = [R_1, S_1, T_1, P_1]^T$ and $S_Y^0 = [R_1, T_1, S_1, P_1]^T$, $i \in \{1, 2\}$. The expected utility functions of players are as follows:

\[ u_X^0(p, q) = v \cdot S_X^0, \quad u_Y^0(p, q) = v \cdot S_Y^0, \quad i \in \{1, 2\}. \]

Denote $G_1 = \{P, \Omega, u, \omega_1\}$, and $G_2 = \{P, \Omega, u, \omega_2\}$, where $P = \{X, Y\}$, $\Omega = \Omega_X \times \Omega_Y$, and $u = \{u_X^0, u_Y^0\}$, $i \in \{1, 2\}$. Thus, in the view of player $Y$, they are playing game $G_1$. In the view of player $X$, they are playing game $G_2$ but player $X$ knows that player $Y$’s cognition is $G_1$. $G_1$ and $G_2$ are shown in Table 2.

**Table 2. Utility matrices in IPD games with misperception.**

|        | (a) $G_1$               | (b) $G_2$               |
|--------|-------------------------|-------------------------|
|        | $Y$                     | $Y$                     |
| $X$    | (c) $(R_1, R_1)$        | (c) $(R_2, R_2)$        |
| $d$    | (d) $(S_1, T_1)$        | (d) $(S_2, T_2)$        |

Let $p_0 = [1, 1, 0, 0]^T$. For $i \in \{1, 2\}$, $p = \alpha S_X^0 + \beta S_Y^0 + \gamma I + p_0$, where $\alpha, \beta, \gamma \in \mathbb{R}$, is called a ZD strategy\textsuperscript{7} of player $X$ in $G_i$ since the strategy makes the two players’ expected utilities subjected to a linear relation:

\[ \alpha u_X^0(p, q) + \beta u_Y^0(p, q) + \gamma = 0, \]

for any player $Y$’s strategy $q$. All available ZD strategies for player $X$ in $G$ can be expressed as $\Xi(\omega) = \{p \in \Omega_X | p = \alpha S_X^0 + \beta S_Y^0 + \gamma I + p_0, \alpha, \beta, \gamma \in \mathbb{R}\}$. Also, the three special ZD strategies are denoted as:

1. (1) equalizer strategy\textsuperscript{7,12}: $p = \beta S_Y^0 + \gamma I + p_0$;
2. (2) extortion strategy\textsuperscript{7,9}: $p = \phi([S_Y^0 - P_1] - \chi(S_Y^0 - P_1)] + p_0, X \geq 1$;
3. (3) generous strategy\textsuperscript{13}: $p = \phi([S_X^0 - R_1] - \chi(S_X^0 - R_1)] + p_0, X \geq 1$.

Here, player $X$ wishes to choose a beneficial strategy to avoid player $Y$’s suspicion of misperception, because player $Y$ believes that player $X$ makes ZD strategies, which can be verified by their strategy sequences. Therefore, player $X$ needs to choose a ZD strategy according to $G_1$ in order to fit player $Y$’s anticipation. To sum up, in our formulation,
• player Y thinks that they are playing game $G_1$, and player X thinks that they are playing game $G_2$;
• player X knows that player Y’s cognition is $G_1$;
• player Y believes that player X chooses ZD strategies;
• player X tends to choose a ZD strategy according to $G_1$ to avoid player Y’s suspicion of misperception.

A beneficial strategy for player X is able to maintain a linear relationship between players’ utilities or improve the supremum or the infimum of its utility in its own cognition. In the following, we aim to analyze player X’s implementation of a ZD strategy in IPD with misperception, and proofs are given in the Supplementary Information.

**Invariance of ZD strategy**

Player X’s ZD strategies may be kept in IPD games with misperception from implementation errors or observation errors. In particular, player X keeps choosing a ZD strategy $p$ in the game $G_1$ in order to avoid player Y’s suspicion about possible misperception. In the view of player X, it can also enforce players’ expected utilities subjected to a linear relationship if $p$ is also a ZD strategy in $G_2$. The following theorem provides a necessary and sufficient condition for the invariance of the linear relationship between players’ utilities.

**Theorem 1** Any ZD strategy $p$ of player X in $G_1$ is also a ZD strategy in $G_2$ if and only if

$$
\frac{R_1 - P_1}{2R_1 - S_1 - T_1} = \frac{R_2 - P_2}{2R_2 - S_2 - T_2}.
$$

If (1) holds, player X can ignore the misperception and choose an arbitrary ZD strategy based on its opponent’s anticipation since it also leads to a linear relationship between players’ utilities, as shown in Figure 1: otherwise, player X can not unscrupulously choose ZD strategies based on player Y’s cognition. There is a player X’s ZD strategy in player Y’s cognition which is not the ZD strategy in player X’s cognition. Further, because of the symmetry of $\omega_1$ and $\omega_2$, player X’s any available ZD strategy $p$ in $G_2$ is also a ZD strategy in $G_1$ if and only if (1) holds. It indicates that $\Xi(\omega_1) = \Xi(\omega_2)$ and player X can choose any ZD strategy based on its own cognition, which does not cause suspicion of the opponent since it is also consistent with player Y’s anticipation. Additionally, the slopes of linear relations between players’ utilities may be different, as also shown in Figure 1, and player X can benefit from the misperception by choosing a ZD strategy to improve the corresponding slope.

In fact, (1) covers the following two cases:

1. $2P_1 = T_i + S_i$, $i \in \{1, 2\}$, is a sufficient condition of (1). Thus, when $2P_1 = T_i + S_i$, $i \in \{1, 2\}$, player X’s any ZD strategy $p$ in $G_1$ is also a ZD strategy in $G_2$. Actually, $2P_1 = T_i + S_i$, $i \in \{1, 2\}$, means that the sum of players’ utilities when players mutual defect is equal to it when only one player chooses defective strategies.

2. \( R_i + P_i = T_i + S_i \), $i \in \{1, 2\}$, is another sufficient condition of (1). Thus, when $R_i + P_i = T_i + S_i$, $i \in \{1, 2\}$, player X’s any ZD strategy $p$ in $G_1$ is also a ZD strategy in $G_2$. Actually, $R_i + P_i = T_i + S_i$, $i \in \{1, 2\}$, means that the game has a balanced structure in utilities\(^{30}\). At this point, the relationship between cooperation rate and efficiency is monotonous, i.e., the higher the cooperation rate of both sides, the greater the efficiency (the sum of players’ utilities).

Furthermore, for the three special ZD strategies, player X can also maintain a linear relationship between players’ utilities in the IPD game with misperception.

**Equalizer strategy**

By choosing equalizer strategies according to player Y’s cognition, player X can unilaterally set player Y’s utilities, as shown in the following corollary.

**Corollary 1** Player X’s any equalizer strategy $p$ in $G_1$ is also an equalizer strategy in $G_2$ if and only if

$$
\frac{R_1 - P_1}{R_2 - P_2} = \frac{R_1 - T_1}{R_2 - T_2} = \frac{R_1 - S_1}{R_2 - S_2}.
$$

(2) is also a sufficient condition of (1). If (2) holds, player X can unilaterally set player Y’s utility by choosing any equalizer strategy according to player Y’s cognition even though they have different cognitions; otherwise, player X can not unscrupulously choose an equalizer strategy based on player Y’s cognition since it may not be an equalizer strategy in player X’s cognition.
Thus, player $X$ randomly generate 100 player $Y$'s strategies, and blue circles are $(u^X_\theta, u^Y_\theta)$, correspondingly. Notice that blue circles are indeed on a cyan line in both (a) and (b).

**Extortion strategy**

By choosing extortion strategies according to player $Y$’s cognition, player $X$ can get an extortionate share, as shown in the following corollary.

**Corollary 2** For player $X$’s extortion strategy $\mathbf{p}$ with extortion factor $\chi > 1$ in $G_1$, $\mathbf{p}$ is also an extortion strategy in $G_2$ if (1) and the following inequality hold:

$$(S_1 - P_1)(R_2 - P_2) - (R_1 - P_1)(T_2 - P_2) - \chi((T_1 - P_1)(R_2 - P_2) - (R_1 - P_1)(T_2 - P_2)) < 0.\quad (3)$$

Player $X$’s extortion strategy in $G_1$, whose extortion factor $\chi$ satisfies (3), can also ensure that player $X$’s utility is not lower than the opponent’s utility in its own cognition. Thus, player $X$ chooses a strategy that satisfies (3), and can also enforce an extortionate share even if there exists misperception.

**Generous strategy**

By choosing generous strategies according to player $Y$’s cognition, player $X$ may also dominate in the game, as reported in the following corollary.

**Corollary 3** For player $X$’s generous strategy $\mathbf{p}$ with generous factor $\chi > 1$ in $G_1$, $\mathbf{p}$ is also a generous strategy in $G_2$ if (1) and the following inequality hold:

$$(S_1 - R_1)(R_2 - P_2) - (R_1 - P_1)(T_2 - R_2) - \chi((T_1 - R_1)(R_2 - P_2) - (R_1 - P_1)(T_2 - R_2)) < 0.\quad (4)$$

A generous strategy ensures that the utility of the player with generous strategies is not higher than the opponent’s utility, but the player dominants in evolving games\textsuperscript{13}. Thus, player $X$’s generous strategy, whose generous factor $\chi$ satisfies (4) based on $Y$’s anticipation, can also dominate in the game in player $X$’s cognition. It is rational for player $X$ to choose generous strategies which satisfy (4) since the misperception does not change their dominant positions.

**Deviation from misperception**

The misperception can lead to a bounded deviation from a linear relationship between players’ expected utilities in player $X$’s cognition. Actually, player $X$ chooses ZD strategies according to player $Y$’s cognition to avoid suspicion, but player $X$ may not enforce a linear relationship between players’ expected utilities in its own cognition. The deviation of the utilities’ relationship is helpful for the player to implement strategies. On the one hand, players’ utilities with misperception go with a bounded deviation from a linear relationship in player $X$’s cognition. Let $\theta$ be the nonzero canonical angles\textsuperscript{31} between the two available ZD strategy sets of $G_1$ and $G_2$, as shown in Figure 2, and we get the following theorem.
focuses on the deviation from an existent linear relationship in player 

Theorem 2 For any player X’s ZD strategy \( \mathbf{p} = \alpha S_X^0 + \beta S_Y^0 + \gamma \mathbf{1} + \mathbf{p}_0 \) in \( G_1 \), there is \( \alpha', \beta', \gamma' \) such that

\[
|\alpha' u_X^{0\gamma}(\mathbf{p}, \mathbf{q}) + \beta' u_Y^{0\gamma}(\mathbf{p}, \mathbf{q}) + \gamma'| \leq \frac{||L_2||_{\infty} \sin \theta, \forall \mathbf{q}},
\]

where \( || \cdot ||_2 \) is the \( l_2 \) norm, \( || \cdot ||_{\infty} \) is the \( l_{\infty} \) norm, and

\[
\theta = \arccos \frac{L_i^T L_i}{||L_i||_2^2}, \quad L_i = [2P_i - S_i - T_i, R_i - P_i, R_i - P_i, T_i + S_i - 2R_i]^T, \quad i \in \{1, 2\}.
\]

Misperception makes players’ utilities a bounded deviation from a linear relationship in player X’s cognition, that is, \( \alpha' u_X + \beta' u_Y + \gamma' = 0 \), even though it is not maintained by choosing ZD strategies in \( G_1 \), as shown in Figure 3(a). By recognizing the difference between \( \omega_1 \) and \( \omega_2 \), player X is able to calculate bounds of players’ utility deviation from misperception.

On the other hand, for a given strategy, the deviation from the corresponding linear relationship is also important, while Theorem 2 focuses on the deviation from an existent linear relationship in player X’s cognition. The misperception can also bring players’ utilities a bounded deviation from the corresponding linear relationship of the ZD strategy in player X’s cognition.

Theorem 3 For player X’s ZD strategy \( \mathbf{p} = \alpha S_X^0 + \beta S_Y^0 + \gamma \mathbf{1} + \mathbf{p}_0 \) in \( G_1 \), the following inequality holds in \( G_2 \).

\[
\min(\Gamma) \leq \alpha u_X^{0\gamma}(\mathbf{p}, \mathbf{q}) + \beta u_Y^{0\gamma}(\mathbf{p}, \mathbf{q}) + \gamma \leq \max(\Gamma),
\]

where

\[
\Gamma = \{(\alpha + \beta)(R_2 - R_1), \alpha(S_2 - S_1) + \beta(T_2 - T_1), \alpha(T_2 - T_1) + \beta(S_2 - S_1), (\alpha + \beta)(P_2 - P_1)\}.
\]

Any ZD strategy of player X based on player Y’s cognition can also enforce players’ utilities subjected to a bounded deviation from the corresponding linear relationship in player X’s cognition, as shown in Figure 3(b). With a ZD strategy \( \mathbf{p} = \alpha S_X^0 + \beta S_Y^0 + \gamma \mathbf{1} + \mathbf{p}_0 \), player X enforces a linear relationship in \( G_1 \), i.e., \( \alpha u_X^{0\gamma}(\mathbf{p}, \mathbf{q}) + \beta u_Y^{0\gamma}(\mathbf{p}, \mathbf{q}) + \gamma = 0 \). Since players’ utilities are \( u_X^{0\gamma} \) and \( u_Y^{0\gamma} \) in \( G_2 \), \( (u_X^{0\gamma}, u_Y^{0\gamma}) \) has a bounded deviation from the corresponding relationship \( \alpha u_X^{0\gamma}(\mathbf{p}, \mathbf{q}) + \beta u_Y^{0\gamma}(\mathbf{p}, \mathbf{q}) + \gamma \).

Benefit from misperception

Player X is able to take advantage of the misperception since it knows player Y’s cognition. To be specific, in IPD without misperception, for any fixed player X’s ZD strategy, its utility is influenced by the opponent’s strategy and is always in a closed interval. Player X can benefit from the misperception by choosing the strategy, which increases the supremum or the infimum of its own utility in IPD with misperception. Besides, for the three special ZD strategies, player X’s ability to improve the supremum/infimum of its own expected utility is shown in Figure 4, and the following results show how player X chooses beneficial strategies.
In this case, by choosing the extortion strategy which satisfies (7), i.e.,

\[ a_i^1 \frac{\gamma}{\beta} > b_i^1, i \in \{1, 2\}, \]

where \( a_i^1 \) and \( b_i^1, i \in \{1, 2\} \) are parameters shown in “Notations”.

Actually, when player \( Y \) chooses the always cooperate (ALLC) strategy \( q = [1, 1, 1, 1]^T \), player \( X \) gets the supremum of the expected utility in \( G_1 \) and player \( X \)'s utility is improved in the IPD game with misperception.

**Extortion strategy**

By choosing extortion strategies according to player \( Y \)'s cognition, player \( X \) can also improve the supremum of its expected utility.

**Corollary 5** For player \( X \)'s extortion strategy \( p \) with extortion factor \( \chi > 1 \) in \( G_1 \), the supremum of player \( X \)'s expected utility in \( G_2 \) is higher than it in \( G_1 \) if

\[ a_i^2 \chi^2 + b_i^2 \chi + c_i^2 < 0, i \in \{1, 2\}, \]

where \( a_i^2, b_i^2, \) and \( c_i^2, i \in \{1, 2\} \) are parameters shown in “Notations”.

If player \( Y \) aims to maximize its own utility with great eagerness, player \( Y \) chooses the ALLC strategy when player \( X \) chooses extortion strategies\(^7\). In this case, by choosing the extortion strategy which satisfies (8), player \( X \) gets the supremum of the expected utility in \( G_1 \), where player \( X \)'s utility is improved in the IPD game with misperception.

**Generous strategy**

By choosing generous strategies according to player \( Y \)'s cognition, player \( X \) can also improve the infimum of its expected utility.

**Corollary 6** For player \( X \)'s generous strategy \( p \) where \( \chi > 1 \), the infimum of player \( X \)'s expected utility in \( G_2 \) is higher than it in \( G_1 \) if

\[ a_i^3 \chi^3 + b_i^3 \chi + c_i^3 < 0, i \in \{1, 2\}, \]

where \( a_i^3, b_i^3, \) and \( c_i^3, i \in \{1, 2\} \) are parameters shown in “Notations”.

---

**Figure 3.** Bounds of Theorems 2 and 3 cover players’ utilities with misperception in player \( X \)'s cognition. Consider \( \omega_1 = [T_1, R_1, P_1, S_1] = [5, 3, 1, 0] \) and \( \omega_2 = [T_2, R_2, P_2, S_2] = [6, \frac{11}{2}, \frac{3}{2}, 0] \). Choose \( p = p_0 S_X^0 + p_0 S_Y^0 + \gamma 1 + p_0 \), where \( (\alpha, \beta, \gamma) = (\frac{1}{30}, -1, \frac{1}{2}) \), and \( (\alpha', \beta', \gamma') = (\frac{38}{165}, -\frac{94}{165}, \frac{151}{165}) \). The red lines denote the relationship between players’ utilities in \( G_1 \). The green lines denote the bounds according to Theorem 2 and Theorem 3, receptively. Then we randomly generate 200 player \( Y \)'s strategies, and the blue circles are \( (u_X^0, u_Y^0) \), correspondingly.
In fact, players may change the parameters and utilities of IPD in others’ cognition, indicating utilities’ relationships when player X chooses an equalizer strategy, an extortion strategy, and a generous strategy in G1, respectively; The yellow area contains all possible relationships between players’ utilities in G2 if player X does not change its strategy. In (a) and (b), r is the supremum of player X’s utility in G1, and r′ is lower than the supremum of player X’s utility in G2; In (c), l is the infimum of player X’s utility in G1, and l′ is lower than the infimum of player X’s utility in G2.

When player X chooses generous strategies, player Y may choose the always defect (ALLD) strategy, i.e., $q = [0, 0, 0, 0]^T$, which is the worst situation for player X since it gets the minimum expected utility in G1. In this case, player X is able to improve its expected utility in the worst situation.

**Discussion**

This paper concentrated on how misperception affects ZD strategies in IPD games. In our problem, player Y is unaware of the different cognitions, but it believes that player X takes a ZD strategy, while player X can detect the misperception. Since each player observes the strategy in sequence, to avoid player Y’s suspicion, player X needs to keep its ZD strategies. Therefore, we explored the ZD strategies in IPD with misperception—a linear relationship between the two players’ expected utilities. In fact, under this affine constraint, player X can ignore the misperception and choose ZD strategies freely. Specifically, we studied the three typical ZD strategies—equalizer, extortion, and generous ones, and moreover, investigated the players’ expected utility deviation from misperception in player X’s cognition. For clarification, we described the deviation not only from the corresponding linear relationship of the ZD strategy but also from another linear relationship that is not directly obtained by player X. Finally, we revealed that the player equipped with ZD strategies may benefit from misperception to improve its own utility. Thus, player X can adopt special equalizer, extortion, or generous strategies to promote the supremum/infimum of its utility in IPD with misperception.

Although both Figures 3(a) and 3(b) illustrates the players’ utility deviation, they are actually derived from different perspectives. Figure 3(a) describes the deviation from a linear relationship, that is, $\alpha' u_X + \beta' u_Y + \gamma' = 0$, where the specific values of $\alpha', \beta', \gamma'$ are not given in Theorem 2. It is helpful for player X to choose beneficial strategies if aiming to get as close to a linear relationship as possible, but no caring about what the linear relationship is. On the other hand, Figure 3(b) indicates that the deviation is derived from a certation linear relation, that is, $\alpha u_X + \beta u_Y + \gamma = 0$, where $\alpha, \beta, \gamma$ are decided by the given ZD strategy. The deviation bounds, according to Theorem 3, are parallel to the linear relationship of the ZD strategy, which helps us analyze the supremum/infimum of player X’s utility with misperception.

Moreover, players may actively utilize misperception to deceive their opponents. For example, players may be able to control their opponents’ observation by interfering with private monitoring, or deliberately mislead their opponents with imitative strategies such as “fake news”. In fact, players may change the parameters and utilities of IPD in others’ cognition by deceiving their opponents. Hence, how the player who adopts ZD strategies benefits from deception in IPD without the opponent’s awareness is also worth analyzing.

**Notations**

$M = [M_{jk}]_{4 \times 4}$ denotes the probability from the last state $k \in \{cc, cd, dc, dd\}$ to the next state $j \in \{cc, cd, dc, dd\}$ in each round, as shown in the following:
\[
M = \begin{bmatrix}
    p_{cc}q_{cc} & p_{cc}(1-q_{cc}) & (1-p_{cc})q_{cc} & (1-p_{cc})(1-q_{cc}) \\
    p_{cd}q_{dc} & p_{cd}(1-q_{dc}) & (1-p_{cd})q_{dc} & (1-p_{cd})(1-q_{dc}) \\
    p_{dc}q_{cd} & p_{dc}(1-q_{cd}) & (1-p_{dc})q_{cd} & (1-p_{dc})(1-q_{cd}) \\
    p_{dd}q_{dd} & p_{dd}(1-q_{dd}) & (1-p_{dd})q_{dd} & (1-p_{dd})(1-q_{dd}) 
\end{bmatrix}.
\]

Denote \( \Upsilon(a,b) = \det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \), \( \Lambda(a,b,c,d) = \det \begin{pmatrix} a_1 & b_2 \\ c_1 & d_2 \end{pmatrix} \), and \( \delta = \max\{|R_2 - R_1|, |S_2 - S_1|, |T_2 - T_1|, |P_2 - P_1|\} \).

The notations in Corollary 4 are shown as follows:
\[
\begin{aligned}
a_1^1 &= \Upsilon(R-S,T-R), \\
b_1^1 &= \Upsilon(R(T-S),R-S) + (R_1 - S_1)(T_2 - R_2)\delta, \\
a_1^2 &= \Lambda(R-S,T-R,T-R,R-S), \\
b_1^2 &= \Lambda(R(T-S),R(T-S),R-S,T-R) + (R_1 - S_1)(R_2 - S_2)\delta.
\end{aligned}
\]

The notations in Corollary 5 are shown as follows:
\[
\begin{aligned}
a_2^1 &= \Upsilon(R(T-S),R-S) - \Upsilon(P(T-R),R-S) + \delta(T_2 - R_2)(R_1 - S_1), \\
b_2^1 &= \Upsilon(R(T-S),R-S) - \Upsilon(P(T-R),T+S-2R) + \delta(T_2 - R_2)(T_1 - S_1), \\
c_2^1 &= (P_1 - P_2 + \delta)(T_2 - R_2)(T_1 - T_1), \\
b_2^2 &= \Lambda(R(T-S),R(T-S),R-S,T-R) - \Lambda(P(T-R),P(R-S),R-S,T-R) + \delta(R_2 - S_2)(R_1 - S_1), \\
b_2^3 &= \Lambda(R(T-S),R(T-S),T-R,R-S) - \Lambda(P(T-R),P(R-S),2R-T-S,T+S-2R) + \delta(R_2 - S_2)(T_1 - S_1), \\
c_2^3 &= (P_1 - P_2 + \delta)(R_2 - S_2)(T_1 - T_1).
\end{aligned}
\]

The notations in Corollary 6 are shown as follows:
\[
\begin{aligned}
a_3^1 &= \Upsilon(P(T-S),T-P) - \Upsilon(R(P-S),T-P) + \delta(T_1 - P_1)(P_2 - S_2), \\
b_3^1 &= \Upsilon(P(T-S),P-S) - \Upsilon(R(P-S),2P-T-S) + \delta(T_1 - S_1)(P_2 - S_2), \\
c_3^1 &= (R_1 - R_2 + \delta)(T_1 - S_1)(P_2 - S_2), \\
a_3^2 &= \Lambda(P(T-S),P(T-S),T-P,P-S) + \Lambda(R(P-S),R(T-P),T-P,P-S) + \delta(T_1 - P_1)(T_2 - P_2), \\
b_3^2 &= \Lambda(P(T-S),P(T-S),P-S,T-P) - \Lambda(R(P-S),R(T-P),2P-T-S,T+S-2P) + \delta(T_1 - S_2)(T_2 - P_2), \\
c_3^2 &= (R_1 - R_2 + \delta)(P_1 - S_1)(T_2 - P_2).
\end{aligned}
\]

References

1. Mailath, G. J. & Samuelson, L. Repeated Games and Reputations: Long-Run Relationships (Oxford University Press, 2006).
2. de Melo, C. M. & Terada, K. The interplay of emotion expressions and strategy in promoting cooperation in the iterated prisoner’s dilemma. Sci. Reports 10, 1–8 (2020).
3. Hilbe, C., Chatterjee, K. & Nowak, M. A. Partners and rivals in direct reciprocity. Nat. Hum. Behav. 2, 469–477 (2018).
4. Wang, Z., Zhou, Y., Lien, J. W., Zheng, J. & Xu, B. Extortion can outperform generosity in the iterated prisoner’s dilemma. Nat. Commun. 7, 1–7 (2016).
5. Glymatsi, N. E. & Knight, V. A. Using a theory of mind to find best responses to memory-one strategies. Sci. Reports 10, 1–9 (2020).
6. Baek, S. K., Jeong, H.-C., Hilbe, C. & Nowak, M. A. Comparing reactive and memory-one strategies of direct reciprocity. Sci. Reports 6, 1–13 (2016).
7. Press, W. H. & Dyson, F. J. Iterated prisoner’s dilemma contains strategies that dominate any evolutionary opponent. Proc. Natl. Acad. Sci. 109, 10409–10413 (2012).
8. Murase, Y. & Baek, S. K. Five rules for friendly rivalry in direct reciprocity. Sci. Reports 10, 1–9 (2020).
9. Beck, L. & Milinski, M. Extortion strategies resist disciplining when higher competitiveness is rewarded with extra gain. Nat. Commun. 10, 1–9 (2019).
10. Govaert, A. & Cao, M. Zero-determinant strategies in repeated multiplayer social dilemmas with discounted payoffs. IEEE Transactions on Autom. Control. (2020).
11. Chen, F., Wu, T. & Wang, L. Evolutionary dynamics of zero-determinant strategies in repeated multiplayer games. arXiv preprint arXiv:2109.06405 (2021).
12. Pan, L., Hao, D., Rong, Z. & Zhou, T. Zero-determinant strategies in iterated public goods game. *Sci. Reports* **5**, 1–10 (2015).
13. Stewart, A. J. & Plotkin, J. B. From extortion to generosity, evolution in the iterated prisoner’s dilemma. *Proc. Natl. Acad. Sci.* **110**, 15348–15353 (2013).
14. Delton, A. W., Krasnow, M. M., Cosmides, L. & Tooby, J. Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proc. Natl. Acad. Sci.* **108**, 13335–13340 (2011).
15. Chen, G., Ming, Y., Hong, Y. & Yi, P. Distributed algorithm for $\varepsilon$-generalized Nash equilibria with uncertain coupled constraints. *Automatica* **123**, 109313 (2021).
16. Solan, E. & Vieille, N. Stochastic games. *Proc. Natl. Acad. Sci.* **112**, 13743–13746 (2015).
17. Cheng, Z., Chen, G. & Hong, Y. Single-leader-multiple-followers Stackelberg security game with hypergame framework. *arXiv preprint arXiv:2107.14625* (2021).
18. Chen, J. & Zhu, Q. Interdependent strategic security risk management with bounded rationality in the internet of things. *IEEE Transactions on Inf. Forensics Secur.* **14**, 2958–2971 (2019).
19. Schlenker, A. et al. Deceiving cyber adversaries: A game theoretic approach. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems*, 892–900 (2018).
20. Rand, D. G., Fudenberg, D. & Dreber, A. It’s the thought that counts: The role of intentions in noisy repeated games. *J. Econ. Behav. & Organ.* **116**, 481–499 (2015).
21. Wang, Z. et al. Exploiting a cognitive bias promotes cooperation in social dilemma experiments. *Nat. Commun.* **9**, 1–7 (2018).
22. Hao, D., Rong, Z. & Zhou, T. Extortion under uncertainty: zero-determinant strategies in noisy games. *Phys. Rev. E* **91**, 052803 (2015).
23. Mamiya, A. & Ichinose, G. Zero-determinant strategies under observation errors in repeated games. *Phys. Rev. E* **102**, 032115 (2020).
24. Kulkarni, A. N., Luo, H., Leslie, N. O., Kamhoua, C. A. & Fu, J. Deceptive labeling: hypergames on graphs for stealthy deception. *IEEE Control. Syst. Lett.* **5**, 977–982 (2020).
25. Heller, Y. & Mohlin, E. Coevolution of deception and preferences: Darwin and Nash meet Machiavelli. *Games Econ. Behav.* **113**, 223–247 (2019).
26. Cranford, E. A. et al. Toward personalized deceptive signaling for cyber defense using cognitive models. *Top. Cogn. Sci.* **12**, 992–1011 (2020).
27. Ortmann, A. & Hertwig, R. The costs of deception: Evidence from psychology. *Exp. Econ.* **5**, 111–131 (2002).
28. Taha, M. A. & Ghoneim, A. Zero-determinant strategies in repeated asymmetric games. *Appl. Math. Comput.* **369**, 124862 (2020).
29. Han, T. A., Perret, C. & Powers, S. T. When to (or not to) trust intelligent machines: Insights from an evolutionary game theory analysis of trust in repeated games. *Cogn. Syst. Res.* **68**, 111–124 (2021).
30. Krueger, J. I. From social projection to social behaviour. *Eur. Rev. Soc. Psychol.* **18**, 1–35 (2008).
31. Qiu, L., Zhang, Y. & Li, C.-K. Unitarily invariant metrics on the Grassmann space. *SIAM J. on Matrix Analysis Appl.* **27**, 507–531 (2005).
32. Imhof, L. A., Fudenberg, D. & Nowak, M. A. Evolutionary cycles of cooperation and defection. *Proc. Natl. Acad. Sci.* **102**, 10797–10800 (2005).
33. Hoang, D. T. et al. Applications of repeated games in wireless networks: A survey. *IEEE Commun. Surv. & Tutorials* **17**, 2102–2135 (2015).
34. Kopp, C., Korb, K. B. & Mills, B. I. Information-theoretic models of deception: Modelling cooperation and diffusion in populations exposed to “fake news”. *PLoS ONE* **13**, e0207383 (2018).
35. Nguyen, T. & Xu, H. Imitative attacker deception in Stackelberg security games. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, 528–534 (2019).

**Acknowledgements**

This work is supported by Shanghai Municipal Science and Technology Major Project under Grant 2021SHZDZX0100 and the National Natural Science Foundation of China under Grants 62173250 and 61733018.
Author contributions
Z.C. conducted experiments. Z.C., G.C, and Y.H. designed research, performed research, and wrote the paper.

Additional information
Supplementary information accompanies this paper at http://www.nature.com/srep.
Competing interests: The authors declare no competing interests.
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SI.pdf