



OPEN

## Misperception influence on zero-determinant strategies in iterated Prisoner's Dilemma

Zhaoyang Cheng<sup>1,2</sup>, Guanpu Chen<sup>1,3</sup> & Yiguang Hong<sup>1,4</sup>✉

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 players<sup>1–3</sup>. 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  $2R > T + S$ <sup>4–6</sup>. 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 quite complicated since players may promote cooperation through past actions. Fortunately, Press and Dyson<sup>7</sup> proposed zero-determinant (ZD) strategies, where the player equipped with ZD strategies can unilaterally enforce the two players' expected utilities subjected to a linear relation. Afterward, various ZD strategies were widely studied in public goods game (PGG), human-computer interaction (HCI), and moving target defense (MTD) problems<sup>8–11</sup>. For example, the equalizer strategy<sup>7,12</sup> is a special ZD strategy that can unilaterally set the opponent's utility. Besides, the player who adopts extortion strategies<sup>7,13</sup> can make that its utility is not lower than the opponent's utility. Conversely, the generous strategy<sup>14,15</sup> 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 interactions<sup>16</sup>, and there have been many models to describe uncertain circumstances in game theory, such as robust games, stochastic games, and hypergames<sup>17–19</sup>. 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 community<sup>20</sup>, and in cyber security problems, hackers may have a confused cognition of the system's TCP/IP stack, which is known to the network administrator<sup>21</sup>. 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 system<sup>22,23</sup>, while players may misunderstand their opponents' strategies, such as some companies relying on private monitoring instead of their opponents' real actions<sup>24,25</sup>.

In fact, the condition for players to trust their cognition is crucial in games with misperception<sup>26,27</sup>. 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<sup>28</sup>. For instance, in psychological experiments, participants' doubts may affect the sponsor's control<sup>29</sup>. Actually, due to the historical information or knowledge from others, a player may know that its opponent takes some given strategies, and

<sup>1</sup>Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Beijing 100190, China. <sup>2</sup>School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100190, China. <sup>3</sup>JD Explore Academy, Beijing 100176, China. <sup>4</sup>Present address: Department of Control Science and Engineering, Tongji University, Shanghai 201804, China. ✉email: yghong@iss.ac.cn

		Y	
		c	d
X	c	(R, R)	(S, T)
	d	(T, S)	(P, P)

**Table 1.** Utility matrix in PD games.

(a) $G_1$				(b) $G_2$			
		Y				Y	
		c	d			c	d
X	c	( $R_1, R_1$ )	( $S_1, T_1$ )	X	c	( $R_2, R_2$ )	( $S_2, T_2$ )
	d	( $T_1, S_1$ )	( $P_1, P_1$ )		d	( $T_2, S_2$ )	( $P_2, P_2$ )

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

moreover, the awareness of the opponent’s ZD strategies has been widely considered in many IPD games<sup>7,9</sup>. In the case when players prefer ZD strategies in IPD with misperception, a player may doubt its cognition if its opponent does not choose ZD strategies as it expects. Nevertheless, most existent works on ZD strategies in IPD with uncertainties, such as ZD strategies with observation errors<sup>24,25</sup> or implementation errors<sup>22,30</sup>, have paid less attention to strategies that maintain players’ cognition.

Therefore, the motivation of this paper is to analyze how misperception affects a player’s ZD strategy without causing its opponent’s 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<sup>22,23,31</sup>. Due to the misperception, the parameter in the real game changes from  $\omega_1 = [T_1, R_1, P_1, S_1]$  to  $\omega_2 = [T_2, R_2, P_2, S_2]$ , and only player X notices the change. Thus, player Y’s cognition of the parameter is  $\omega_1$ , while player X’s cognition of the parameter is  $\omega_2$ . In each round, player X chooses a strategy from its strategy set  $\Omega_X = \{\mathbf{p} = [p_{cc}, p_{cd}, p_{dc}, p_{dd}]^T | p_{xy} \in [0, 1], xy \in \{cc, cd, dc, dd\}\}$ , e.g.,  $p_{xy}$  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  $\Omega_Y = \{\mathbf{q} = [q_{cc}, q_{dc}, q_{cd}, q_{dd}]^T | q_{xy} \in [0, 1], xy \in \{cc, dc, cd, dd\}\}$ . According to Press and Dyson<sup>7</sup>, this game can be characterized by a Markov chain with a state transition matrix  $M = [M_{jk}]_{4 \times 4}$  (see “Notations” for details). Denote  $\mathbf{v} = [v_{cc}, v_{cd}, v_{dc}, v_{dd}]^T$  as a probability vector such that  $\mathbf{v}^T M = \mathbf{v}^T$  and  $v_{cc} + v_{cd} + v_{dc} + v_{dd} = 1$ . Let  $\mathbf{S}_X^{\omega_i} = [R_i, S_i, T_i, P_i]^T$ , and  $\mathbf{S}_Y^{\omega_i} = [R_i, T_i, S_i, P_i]^T$ ,  $i \in \{1, 2\}$ . The expected utility functions of players are as follows:

$$u_X^{\omega_i}(\mathbf{p}, \mathbf{q}) = \mathbf{v} \cdot \mathbf{S}_X^{\omega_i}, \quad u_Y^{\omega_i}(\mathbf{p}, \mathbf{q}) = \mathbf{v} \cdot \mathbf{S}_Y^{\omega_i}, \quad i \in \{1, 2\}.$$

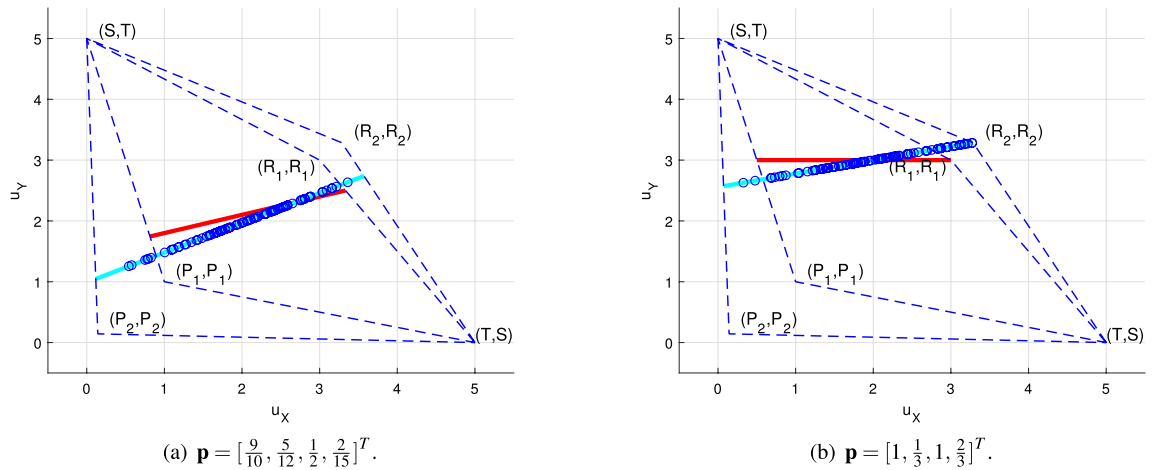
Denote  $G_1 = \{\mathbf{P}, \Omega, \mathbf{u}, \omega_1\}$ , and  $G_2 = \{\mathbf{P}, \Omega, \mathbf{u}, \omega_2\}$ , where  $\mathbf{P} = \{X, Y\}$ ,  $\Omega = \Omega_X \times \Omega_Y$ , and  $\mathbf{u} = \{u_X^{\omega_i}, u_Y^{\omega_i}\}$ ,  $i \in \{1, 2\}$ . Thus, the actual utilities of players are obtained through  $G_2$ , and 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.

Let  $\mathbf{p}_0 = [1, 1, 0, 0]^T$ . For  $i \in \{1, 2\}$ ,  $\mathbf{p} = \alpha \mathbf{S}_X^{\omega_i} + \beta \mathbf{S}_Y^{\omega_i} + \gamma \mathbf{1} + \mathbf{p}_0$ , where  $\alpha, \beta, \gamma \in \mathbb{R}$ , is called a ZD strategy<sup>7</sup> of player X in  $G_i$  since the strategy makes the two players’ expected utilities subjected to a linear relation:

$$\alpha u_X^{\omega_i}(\mathbf{p}, \mathbf{q}) + \beta u_Y^{\omega_i}(\mathbf{p}, \mathbf{q}) + \gamma = 0,$$

for any player Y’s strategy  $\mathbf{q}$ . All available ZD strategies for player X in  $G$  can be expressed as  $\Xi(\omega_i) = \{\mathbf{p} \in \Omega_X | \mathbf{p} = \alpha \mathbf{S}_X^{\omega_i} + \beta \mathbf{S}_Y^{\omega_i} + \gamma \mathbf{1} + \mathbf{p}_0, \alpha, \beta, \gamma \in \mathbb{R}\}$ . Also, the three special ZD strategies are denoted as:

- (1) equalizer strategy<sup>7,12</sup>:  $\mathbf{p} = \beta \mathbf{S}_Y^{\omega_i} + \gamma \mathbf{1} + \mathbf{p}_0$ ;
- (2) extortion strategy<sup>7,13</sup>:  $\mathbf{p} = \phi[(\mathbf{S}_X^{\omega_i} - P_i \mathbf{1}) - \chi(\mathbf{S}_Y^{\omega_i} - P_i \mathbf{1})] + \mathbf{p}_0, \chi \geq 1$ ;
- (3) generous strategy<sup>14,15</sup>:  $\mathbf{p} = \phi[(\mathbf{S}_X^{\omega_i} - R_i \mathbf{1}) - \chi(\mathbf{S}_Y^{\omega_i} - R_i \mathbf{1})] + \mathbf{p}_0, \chi \geq 1$ .



**Figure 1.** Player X can also enforce a linear relationship between players’ utilities in its own cognition. Let  $\omega_1 = [T, R_1, P_1, S] = [5, 3, 1, 0]$  and  $\omega_2 = [T, R_2, P_2, S] = [5, \frac{23}{7}, \frac{1}{7}, 0]$ , which satisfy (1). Consider that player X chooses two different ZD strategies in (a) and (b), respectively, and the red lines describe the relationships between players’ utilities in  $G_1$ . We randomly generate 100 player Y’s strategies, and blue circles are  $(u_X^{\omega_2}, u_Y^{\omega_2})$ , correspondingly. Notice that blue circles are indeed on a cyan line in both (a) and (b).

Based on the past experience, player Y knows that player X prefers ZD strategies, which has been widely considered in many IPD games<sup>7,9</sup>. To avoid that player Y notices the change, which may result in potential decrease of player X’s utility<sup>21</sup> or collapse of the model<sup>28</sup>, player X keeps choosing ZD strategies according to  $G_1$ , such that the strategy sequence matches player Y’s anticipation. To sum up, in our formulation,

- the real game is  $G_2$ ;
- 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.

In fact, player X can benefit from the misperception through the ZD strategy. For example, player X can adopt a generous strategy in  $G_1$  to not only promote player Y’s cooperation behavior, but also make player X’s utility higher than that of player Y, if the generous strategy is an extortion strategy in  $G_2$ . 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  $\mathbf{p}$  in  $G_1$  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  $\mathbf{p}$  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.

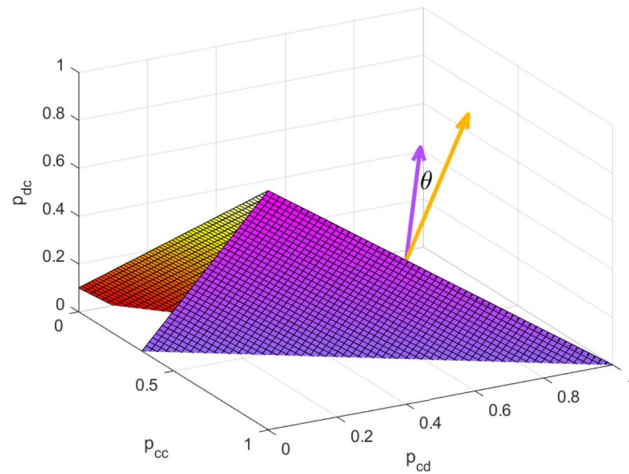
**Theorem 1** Any ZD strategy  $\mathbf{p}$  of player X in  $G_1$  is also a ZD strategy in  $G_2$  if and only if

$$\frac{R_1 - P_1}{2R_1 - S_1 - T_1} = \frac{R_2 - P_2}{2R_2 - S_2 - T_2}. \tag{1}$$

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 Fig. 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  $\mathbf{p}$  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 Fig. 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)  $2P_i = T_i + S_i, i \in \{1, 2\}$ , is a sufficient condition of (1). Thus, when  $2P_i = T_i + S_i, i \in \{1, 2\}$ , player X’s any ZD strategy  $\mathbf{p}$  in  $G_1$  is also a ZD strategy in  $G_2$ . Actually,  $2P_i = T_i + S_i, i \in \{1, 2\}$ , means that the sum of



**Figure 2.** The form of  $\theta$  in the IPD game with misperception. Consider  $\omega_1 = [T_1, R_1, P_1, S_1] = [5, 3, 1, 0]$  and  $\omega_2 = [T_2, R_2, P_2, S_2] = [\frac{13}{2}, 6, 1, 0]$ . Suppose  $p_{dd} = 0$  since it does not influence nonzero canonical angles. The purple (yellow) plane is the available ZD strategy set in  $G_1$  ( $G_2$ ) and the purple (yellow) vector is its normal vector. Clearly,  $\theta$  is the angle between two normal vectors, which is also the nonzero canonical angle between the available ZD strategy set of  $G_1$  and it of  $G_2$ .

- players' utilities when players mutual defect is equal to that 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  $\mathbf{p}$  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<sup>32</sup>. 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  $\mathbf{p}$  in  $G_1$  is also an equalizer strategy in  $G_2$  if and only if

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

(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 in  $G_1$  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.

*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.

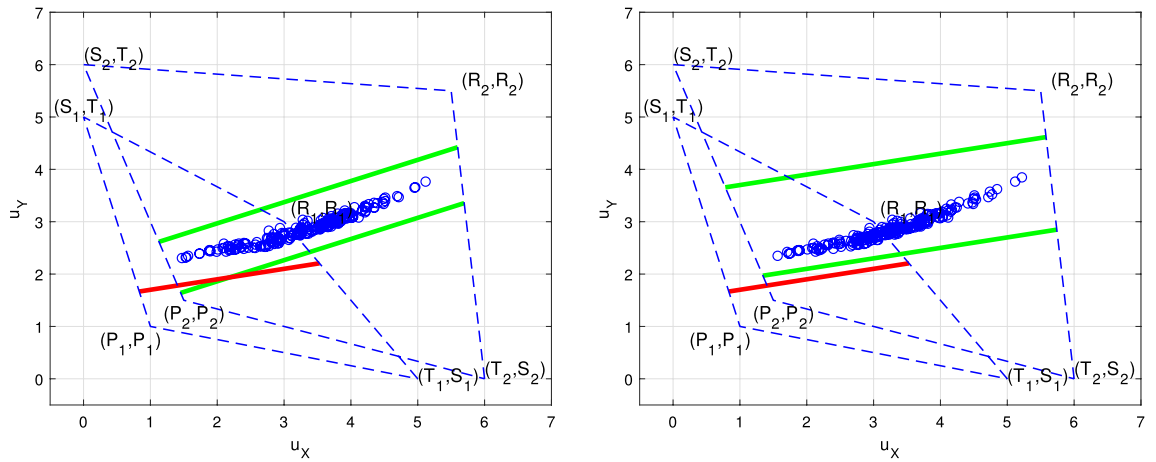
**Corollary 2** For player  $X$ 's extortion strategy  $\mathbf{p}$  with extortion factor  $\chi > 1$  in  $G_1$ ,  $\mathbf{p}$  is also an extortion strategy in  $G_2$  if (1) and the following inequality hold:

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

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:



(a) Deviation from a linear relationship

(b) Deviation from the certain linear relationship

**Figure 3.** The relation between bounds of Theorems 2 and 3 and players' utilities in  $G_2$ . 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  $\mathbf{p} = \alpha \mathbf{S}_X^{\omega_1} + \beta \mathbf{S}_Y^{\omega_1} + \gamma \mathbf{1} + \mathbf{p}_0$ , where  $(\alpha, \beta, \gamma) = (\frac{1}{30}, -\frac{1}{6}, \frac{1}{4})$ , and  $(\alpha', \beta', \gamma') = (\frac{38}{165}, -\frac{94}{165}, \frac{151}{165})$ . The red lines describe the relationship between players' utilities in  $G_1$ . The green lines describe the bounds according to Theorems 2 and 3. Then we randomly generate 200 player Y's strategies, and the blue circles are  $(u_X^{\omega_2}, u_Y^{\omega_2})$ , correspondingly.

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

A generous strategy ensures that the utility of the player with generous strategies is not higher than the opponent's utility, but the player dominates in evolving games<sup>14,33</sup>. 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 a ZD strategy to avoid player Y's 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<sup>34</sup> between the two available ZD strategy sets of  $G_1$  and  $G_2$ , as shown in Fig. 2, and we get the following theorem.

**Theorem 2** For any player X's ZD strategy  $\mathbf{p} = \alpha \mathbf{S}_X^{\omega_1} + \beta \mathbf{S}_Y^{\omega_1} + \gamma \mathbf{1} + \mathbf{p}_0$  in  $G_1$ , there is  $\alpha', \beta', \gamma'$  such that

$$|\alpha' u_X^{\omega_2}(\mathbf{p}, \mathbf{q}) + \beta' u_Y^{\omega_2}(\mathbf{p}, \mathbf{q}) + \gamma'| \leq \|\mathbf{p}\|_2 \frac{\|L_2\|_\infty}{\|L_2\|_2} \sin\theta, \forall \mathbf{q},$$

where  $\|\cdot\|_2$  is the  $l_2$  norm,  $\|\cdot\|_\infty$  is the  $l_\infty$  norm, and

$$\theta = \arccos \frac{L_1^T L_2}{\|L_1\|_2 \|L_2\|_2}, L_i = [2P_i - S_i - T_i, R_i - P_i, R_i - P_i, T_i + S_i - 2R_i]^T, i \in \{1, 2\}.$$

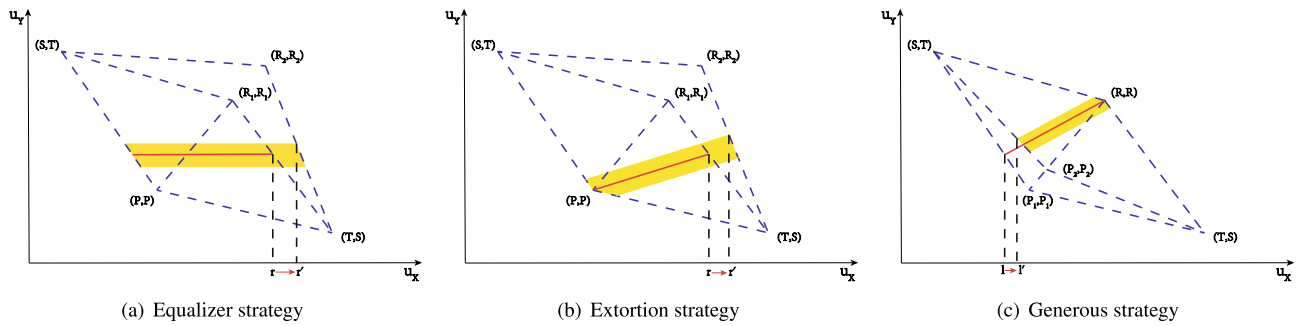
Misperception makes players' utilities a bounded deviation from a linear relationship in player X's cognition, that is,  $\alpha' u_X + \beta' u_Y + \gamma' = 0$ , even though it is not maintained by choosing ZD strategies in  $G_1$ , as shown in Fig. 3a. By recognizing the difference between  $\omega_1$  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  $\mathbf{p} = \alpha \mathbf{S}_X^{\omega_1} + \beta \mathbf{S}_Y^{\omega_1} + \gamma \mathbf{1} + \mathbf{p}_0$  in  $G_1$ , the following inequality holds in  $G_2$ ,

$$\min(\Gamma) \leq \alpha u_X^{\omega_2}(\mathbf{p}, \mathbf{q}) + \beta u_Y^{\omega_2}(\mathbf{p}, \mathbf{q}) + \gamma \leq \max(\Gamma),$$

where



**Figure 4.** Player X can use either equalizer strategies and extortion strategies to raise the supremum of its expected utility or generous strategies to raise the infimum of its expected utility. (a) and (b) consider that  $\omega_1 = [T, R_1, P, S]$  and  $\omega_2 = [T, R_2, P, S]$ , where  $R_1 \neq R_2$ ; (c) considers that  $\omega_1 = [T, R, P_1, S]$  and  $\omega_2 = [T, R, P_2, S]$ , where  $P_1 \neq P_2$ . The red lines in (a), (b), and (c) describe utilities' relationships when player X chooses an equalizer strategy, an extortion strategy, and a generous strategy in  $G_1$ , respectively; The yellow area contains all possible relationships between players' utilities in  $G_2$  if player X does not change its strategy. In (a) and (b),  $r$  is the supremum of player X's utility in  $G_1$ , and  $r'$  is lower than the supremum of player X's utility in  $G_2$ ; In (c),  $l$  is the infimum of player X's utility in  $G_1$ , and  $l'$  is lower than the infimum of player X's utility in  $G_2$ .

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

Any ZD strategy of player X based on player Y's cognition can enforce players' utilities subjected to a bounded deviation from the corresponding linear relationship in player X's cognition, as shown in Fig. 3b. With a ZD strategy  $\mathbf{p} = \alpha S_X^{\omega_1} + \beta S_Y^{\omega_1} + \gamma \mathbf{1} + \mathbf{p}_0$ , player X enforces a linear relationship in  $G_1$ , i.e.,  $\alpha u_X^{\omega_1}(\mathbf{p}, \mathbf{q}) + \beta u_Y^{\omega_1}(\mathbf{p}, \mathbf{q}) + \gamma = 0$ . Since players' utilities are  $u_X^{\omega_2}$  and  $u_Y^{\omega_2}$  in  $G_2$ ,  $(u_X^{\omega_2}, u_Y^{\omega_2})$  has a bounded deviation from the corresponding relationship  $\alpha u_X^{\omega_2}(\mathbf{p}, \mathbf{q}) + \beta u_Y^{\omega_2}(\mathbf{p}, \mathbf{q}) + \gamma$ .

**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 Fig. 4, and the following results show how player X chooses beneficial strategies.

*Equalizer strategy.* By choosing equalizer strategies according to player Y's cognition, player X can improve the supremum of its expected utility.

**Corollary 4** For player X's equalizer strategy  $\mathbf{p} = \beta S_Y^{\omega_1} + \gamma \mathbf{1} + \mathbf{p}_0$ ,  $\beta \neq 0$ , in  $G_1$ , the supremum of player X's expected utility in  $G_2$  is larger than that in  $G_1$ , if

$$a_i^1 \frac{\gamma}{\beta} > b_i^1, i \in \{1, 2\}, \tag{5}$$

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<sup>35</sup>, i.e.,  $\mathbf{q} = [1, 1, 1, 1]^T$ , 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  $\mathbf{p}$  with extortion factor  $\chi > 1$  in  $G_1$ , the supremum of player X's expected utility in  $G_2$  is larger than that in  $G_1$  if

$$a_i^2 \chi^2 + b_i^2 \chi + c_i^2 < 0, i \in \{1, 2\}, \tag{6}$$

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<sup>7</sup>. In this case, by choosing the extortion strategy which satisfies (6), 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  $\mathbf{p}$  where  $\chi > 1$ , the infimum of player  $X$ 's expected utility in  $G_2$  is larger than that in  $G_1$  if

$$a_i^3 \chi^2 + b_i^3 \chi + c_i^3 < 0, i \in \{1, 2\}, \tag{7}$$

where  $a_i^3, b_i^3$ , and  $c_i^3, i \in \{1, 2\}$  are parameters shown in "Notations".

When player  $X$  chooses generous strategies, player  $Y$  may choose the always defect (ALLD) strategy<sup>35</sup>, i.e.,  $\mathbf{q} = [0, 0, 0, 0]^T$ , which is the worst situation for player  $X$  since it gets the minimum expected utility in  $G_1$ . In this case, player  $X$  is able to improve its expected utility in the worst situation.

**Discussion**

This paper concentrates 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 have 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 have studied the three typical ZD strategies—equalizer, extortion, and generous ones, and moreover, we have investigated the players' expected utility deviation from misperception in player  $X$ 's cognition. For clarification, we have 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 have 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 Fig. 3a, b illustrates the players' utility deviation, they are actually derived from different perspectives. Figure 3a 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, Fig. 3b 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 adopt misperception to deceive their opponents. For example, players may be able to control their opponents' observation by interfering with private monitoring<sup>36</sup>, or deliberately mislead their opponents with imitative strategies such as "fake news"<sup>37,38</sup>. 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. Since the ZD strategy has also been widely applied in other complicated situations, such as non-symmetric games<sup>39</sup>, PGG<sup>10</sup>, and evolutionary situations<sup>33</sup>, the misperception influence analysis will be extended to the ZD strategies in these practical fields.

**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:

$$M = \begin{bmatrix} P_{cc}q_{cc} & P_{cc}(1 - q_{cc}) & (1 - P_{cc})q_{cc} & (1 - P_{cc})(1 - q_{cc}) \\ P_{cd}q_{dc} & P_{cd}(1 - q_{dc}) & (1 - P_{cd})q_{dc} & (1 - P_{cd})(1 - q_{dc}) \\ P_{dc}q_{cd} & P_{dc}(1 - q_{cd}) & (1 - P_{dc})q_{cd} & (1 - P_{dc})(1 - q_{cd}) \\ P_{dd}q_{dd} & P_{dd}(1 - q_{dd}) & (1 - P_{dd})q_{dd} & (1 - P_{dd})(1 - q_{dd}) \end{bmatrix}.$$

Thus,  $M$  is regular when all elements of  $M$  are positive, e.g.,  $0 < p_{xy}, q_{xy} < 1, xy \in \{cc, cd, dc, dd\}$ . Denote

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

The notations in Corollary 4 are shown as follows:

$$\begin{aligned} a_1^1 &= \Upsilon(R - S, T - R), \\ b_1^1 &= \Upsilon(R(T - S), R - S) + (R_1 - S_1)(T_2 - R_2)\delta, \\ a_2^1 &= \Lambda(R - S, T - R, T - R, R - S), \\ b_2^1 &= \Lambda(R(T - S), R(T - S), R - S, T - R) + (R_1 - S_1)(R_2 - S_2)\delta. \end{aligned}$$

The notations in Corollary 5 are shown as follows:

$$\begin{aligned} a_1^2 &= \Upsilon(R(T - S), R - S) - \Upsilon(P(T - R), R - S) + \delta(T_2 - R_2)(R_1 - S_1), \\ b_1^2 &= \Upsilon(R(T - S), R - S) - \Upsilon(P(T - R), T + S - 2R) + \delta(T_2 - R_2)(T_1 - S_1), \\ c_1^2 &= (P_1 - P_2 + \delta)(T_2 - R_2)(T_1 - R_1), \\ a_2^2 &= \Lambda(R(T - S), R(T - S), R - S, T - R) - \Lambda(P(T - R), P(R - S), R - S, T - R) + \delta(R_2 - S_2)(R_1 - S_1), \\ b_2^2 &= \Lambda(R(T - S), R(T - S), T - R, R - S) - \Lambda(P(T - R), P(R - S), 2R - T - S, T + S - 2R) + \delta(R_2 - S_2)(T_1 - S_1), \\ c_2^2 &= (P_1 - P_2 + \delta)(R_2 - S_2)(T_1 - R_1). \end{aligned}$$

The notations in Corollary 6 are shown as follows:

$$a_1^3 = \Upsilon(P(T-S), T-P) - \Upsilon(R(P-S), T-P) + \delta(T_1 - P_1)(P_2 - S_2),$$

$$b_1^3 = \Upsilon(P(T-S), P-S) - \Upsilon(R(P-S), 2P-T-S) + \delta(T_1 - S_1)(P_2 - S_2),$$

$$c_1^3 = (R_1 - R_2 + \delta)(T_1 - S_1)(P_2 - S_2),$$

$$a_2^3 = \Lambda(P(T-S), P(T-S), T-P, P-S) - \Lambda(R(P-S), R(T-P), T-P, P-S) + \delta(T_1 - P_1)(T_2 - P_2),$$

$$b_2^3 = \Lambda(P(T-S), P(T-S), P-S, T-P) - \Lambda(R(P-S), R(T-P), 2P-T-S, T+S-2P) + \delta(T_1 - S_2)(T_2 - P_2),$$

$$c_2^3 = (R_1 - R_2 + \delta)(P_1 - S_1)(T_2 - P_2).$$

Received: 13 December 2021; Accepted: 17 February 2022

Published online: 25 March 2022

## References

1. Mailath, G. J. & Samuelson, L. *Repeated games and reputations: long-run relationships* (Oxford University Press, Oxford, 2006).
2. de Melo, C. M. & Terada, K. The interplay of emotion expressions and strategy in promoting cooperation in the iterated Prisoners Dilemma. *Sci. Rep.* **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. Glynatsi, N. E. & Knight, V. A. Using a theory of mind to find best responses to memory-one strategies. *Sci. Rep.* **10**, 1–9 (2020).
5. Baek, S. K., Jeong, H.-C., Hilbe, C. & Nowak, M. A. Comparing reactive and memory-one strategies of direct reciprocity. *Sci. Rep.* **6**, 1–13 (2016).
6. Murase, Y. & Baek, S. K. Five rules for friendly rivalry in direct reciprocity. *Sci. Rep.* **10**, 1–9 (2020).
7. Press, W. H. & Dyson, F. J. Iterated Prisoners Dilemma contains strategies that dominate any evolutionary opponent. *Proc. Natl. Acad. Sci.* **109**, 10409–10413 (2012).
8. Wang, S., Shi, H., Hu, Q., Lin, B. & Cheng, X. Moving target defense for internet of things based on the zero-determinant theory. *IEEE Internet Things J.* **7**, 661–668 (2019).
9. Wang, Z., Zhou, Y., Lien, J. W., Zheng, J. & Xu, B. Extortion can outperform generosity in the iterated Prisoners Dilemma. *Nat. Commun.* **7**, 1–7 (2016).
10. Govaert, A. & Cao, M. Zero-determinant strategies in repeated multiplayer social dilemmas with discounted payoffs. *IEEE Trans. Autom. Control* **66**, 4575–4588 (2021).
11. Chen, F., Wu, T. & Wang, L. Evolutionary dynamics of zero-determinant strategies in repeated multiplayer games. arXiv preprint [arXiv:2109.06405](https://arxiv.org/abs/2109.06405) (2021).
12. Pan, L., Hao, D., Rong, Z. & Zhou, T. Zero-determinant strategies in iterated public goods game. *Sci. Rep.* **5**, 1–10 (2015).
13. Becks, L. & Milinski, M. Extortion strategies resist disciplining when higher competitiveness is rewarded with extra gain. *Nat. Commun.* **10**, 1–9 (2019).
14. Stewart, A. J. & Plotkin, J. B. From extortion to generosity, evolution in the iterated Prisoners Dilemma. *Proc. Natl. Acad. Sci.* **110**, 15348–15353 (2013).
15. Akin, E. The iterated prisoner's dilemma: Good strategies and their dynamics. In *Ergodic Theory*, 77–107 (2016).
16. 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).
17. Chen, G., Ming, Y., Hong, Y. & Yi, P. Distributed algorithm for  $\epsilon$ -generalized Nash equilibria with uncertain coupled constraints. *Automatica* **123**, 109313 (2021).
18. Solan, E. & Vieille, N. Stochastic games. *Proc. Natl. Acad. Sci.* **112**, 13743–13746 (2015).
19. Cheng, Z., Chen, G. & Hong, Y. Single-leader-multiple-followers Stackelberg security game with hypergame framework. *IEEE Trans. Inf. Forensics Secur.* **14**, 954–969. <https://doi.org/10.1109/TIFS.2022.3155294> (2022).
20. Chen, J. & Zhu, Q. Interdependent strategic security risk management with bounded rationality in the internet of things. *IEEE Trans. Inf. Forensics Secur.* **14**, 2958–2971 (2019).
21. Schlenker, A. *et al.* Deceiving cyber adversaries: A game theoretic approach. In: Proceedings of the 17th international conference on autonomous agents and multiagent systems, pp. 892–900 (2018).
22. Rand, D. G., Fudenberg, D. & Dreber, A. Its the thought that counts: The role of intentions in noisy repeated games. *J. Econ. Behav. Organ.* **116**, 481–499 (2015).
23. Wang, Z. *et al.* Exploiting a cognitive bias promotes cooperation in social dilemma experiments. *Nat. Commun.* **9**, 1–7 (2018).
24. Hao, D., Rong, Z. & Zhou, T. Extortion under uncertainty: Zero-determinant strategies in noisy games. *Phys. Rev. E* **91**, 052803 (2015).
25. Mamiya, A. & Ichinose, G. Strategies that enforce linear payoff relationships under observation errors in repeated Prisoners Dilemma game. *J. Theor. Biol.* **477**, 63–76 (2019).
26. 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).
27. Heller, Y. & Mohlin, E. Coevolution of deception and preferences: Darwin and Nash meet Machiavelli. *Games Econom. Behav.* **113**, 223–247 (2019).
28. Cranford, E. A. *et al.* Toward personalized deceptive signaling for cyber defense using cognitive models. *Top. Cogn. Sci.* **12**, 992–1011 (2020).
29. Ortmann, A. & Hertwig, R. The costs of deception: Evidence from psychology. *Exp. Econ.* **5**, 111–131 (2002).
30. Taha, M. A. & Ghoneim, A. Zero-determinant strategies in repeated asymmetric games. *Appl. Math. Comput.* **369**, 124862 (2020).
31. 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).
32. Krueger, J. I. From social projection to social behaviour. *Eur. Rev. Soc. Psychol.* **18**, 1–35 (2008).
33. Hilbe, C., Nowak, M. A. & Sigmund, K. Evolution of extortion in iterated prisoners dilemma games. *Proc. Natl. Acad. Sci.* **110**, 6913–6918 (2013).
34. Qiu, L., Zhang, Y. & Li, C.-K. Unitarily invariant metrics on the Grassmann space. *SIAM J. Matrix Anal. Appl.* **27**, 507–531 (2005).
35. Imhof, L. A., Fudenberg, D. & Nowak, M. A. Evolutionary cycles of cooperation and defection. *Proc. Natl. Acad. Sci.* **102**, 10797–10800 (2005).
36. Hoang, D. T. *et al.* Applications of repeated games in wireless networks: A survey. *IEEE Commun. Surv. Tutor.* **17**, 2102–2135 (2015).
37. 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).



38. 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).
39. Akin, E. What you gotta know to play good in the iterated prisoners dilemma. *Games* **6**, 175–190 (2015).

### 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.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-08750-8>.

**Correspondence** and requests for materials should be addressed to Y.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022