The Robustness of Zero-Determinant Strategies in Iterated Prisoner's Dilemma Games

Jing Chen[*,a] and Aleksey Zinger[b]

[a]*Department of Computer Science, Stony Brook University, Stony Brook, NY 11794, USA*

[b]*Department of Mathematics, Stony Brook University, Stony Brook, NY 11794, USA*

May 16, 2014

Abstract

Press and Dyson (2012) discovered a special set of strategies in two-player Iterated Prisoner's Dilemma games, the *zero-determinant (ZD) strategies*. Surprisingly, a player using such strategies can unilaterally enforce a linear relation between the payoffs of the two players. In particular, with a subclass of such strategies, the *extortionate strategies*, the former player obtains an advantageous share of the total payoff of the players, and the other player's best response is to always cooperate, by doing which he maximizes the payoff of the extortioner as well. When an extortionate player faces a player who is not aware of the theory of ZD strategies and improves his own payoff by adaptively changing his strategy following some unknown dynamics, Press and Dyson conjecture that there always exist adapting paths for the latter leading to the maximum possible scores for both players.

In this work we confirm their conjecture in a very strong sense, not just for extortionate strategies, but for *all* ZD strategies that impose positive correlations between the players' payoffs. We show that not only the conjectured adapting paths always exist, but that actually *every* adapting path leads to the maximum possible scores, although some paths may not lead to the unconditional cooperation by the adapting player. This is true even in the rare cases where the setup of Press and Dyson is not directly applicable. Our result shows that ZD strategies are even more powerful than as pointed out by their discoverers. Given our result, the player using ZD strategies is assured that she will receive the maximum payoff attainable under the desired payoff relation she imposes, without knowing how the other player will evolve. This makes the use of ZD strategies even more desirable for sentient players.

Keywords: Iterated Prisoner's Dilemma, Zero-Determinant strategy, adapting player, adapting path, cooperative behavior.

1 Introduction

The two-player Iterated Prisoner's Dilemma (IPD) game is one of the standard models for studying the emergence of cooperative behavior among competitive players. It has long been investigated in economics, political science, evolutionary biology, and computer science (see [8], [5], [3], [18], [4], [19], [16], [14], [15], [7], [9], [13], [12] and [6], as just a few examples). As IPD has been so widely studied, it was surprising when Press and Dyson [17] discovered a completely new property of this game, namely, the existence of Zero-Determinant (ZD) strategies. Roughly speaking, such strategies allow one player to unilaterally set the payoff score of the other or to enforce a linear relation between the two players' scores, as opposite to the previous general belief that no ultimatum strategy can enforce any specific kind of outcome. Among such strategies, of particular interest

[*]Corresponding author at: Department of Computer Science, Stony Brook University, Stony Brook, NY 11794, USA. Tel.:+1 631 632 1827.

E-mail addresses: jingchen@cs.stonybrook.edu, azinger@math.sunysb.edu

Player X \ *Player Y*	C	D
C	(R,R)	(S,T)
D	(T,S)	(P,P)

Figure 1: Scores for the players X and Y in each of the four outcomes in a single play of Prisoner's Dilemma.

are the so-called extortionate strategies [17], in which the sentient player takes a larger share of the total benefit, and generous strategies [2, 20, 21], in which the sentient player takes a larger share of the total loss from the full-cooperation rewards. The results of [17] have led to completely new viewpoints on IPD. Since then, the properties of ZD strategies, including extortionate strategies in arbitrary IPD games and generous strategies in donation games, have been actively studied; see [20], [10], [21], [1], [2], [22], and [11].

The game under consideration here is of discrete time and with infinitely many rounds. In each round, the same two players, X and Y, play the one-shot Prisoner's Dilemma (PD). As illustrated in Figure 1, each player can choose to cooperate (C) or to defect (D), without knowing the other's choice. If both cooperate, then each receives score R. If both defect, then each receives a smaller score P. If one cooperates and the other defects, then the defector rips off a score T larger than R, and the cooperator gets ripped off with a score of S smaller than P. The literature typically assumes $2R > T + S > 2P$, so that the total score of the players is maximized when both cooperate. For example, $(T, R, P, S) = (5, 3, 1, 0)$ is a conventional realization of the parameters.

Press and Dyson [17] assume that both players have memory of length 1, i.e. what a player does in the current round only depends on the outcome of the previous round, rather than the whole history of the play or the number of rounds played. Accordingly, a (mixed) strategy of a player consists of a mapping from the four possible outcomes of PD to the probabilities of cooperating. The strategies of the two players together with a starting outcome determine a Markov chain. The players' payoff scores, s_X and s_Y, are defined to be the expected scores they would receive under the stationary distribution of the Markov chain.

A ZD strategy of player X guarantees

$$s_X - K = \chi(s_Y - K)$$

for specified values of χ and K satisfying certain conditions, no matter what strategy Y uses. For s_X and s_Y to be positively correlated, one needs $\chi \geq 1$ (the cases $\chi \in (0, 1)$ do not correspond to ZD strategies). Facing such ZD strategies, when Y adjusts his own strategy to increase his score, he increases X's score even more, and when he achieves his own maximum score, X's score is also maximized. Both extortionate and generous strategies are *positively correlated ZD (pcZD) strategies*, with the former satisfying $K = P$ and the latter $K = R$.

As pointed out by [17], for extortionate strategies, the scores of both players are maximized when Y cooperates unconditionally, namely, uses the strategy $(1, 1, 1, 1)$.

An extortionate player facing an adapting player. One question which is not completely answered by [17] and not considered by previous followups is the following. *What should a player X witting of extortionate strategies do if she believes that her opponent Y is an adapting player?* An adapting player is one who tries to improve his own score following some optimization scheme (perhaps known only to him), but without explicitly considering or trying to alter the strategy of X. Such a player is called an evolutionary player in [17], but "evolutionary" already has a specific (and different) meaning in the context of game theory, and thus we use "adapting" instead, to avoid confusion. The answer to the question above depends on how Y adapts. Although it is of Y's best interest to unconditionally cooperate, he may not realize this fact and may only make local movement to gradually improve his score. Since the direction of improvement is not unique, in principle Y might end up at a local optimum and leave X with a score much smaller than what she expects when Y unconditionally cooperates. If this can happen, then X would use an extortionate strategy only if she believes that Y will take a desirable adapting path (roughly speaking, an *adapting path* is a smooth map from time to Y's strategies such that Y's utility increases along time —formally defined in Section 3), and would otherwise continue monitoring the behavior of Y and change her strategy when necessary.

Press and Dyson conjecture that in all cases, that is, with different parameters (R, T, S, P), different starting points of IPD, and different original strategies of Y, there *exist* adapting paths of Y that lead to the globally maximum scores when X applies an extortionate strategy. However, the existence of desirable adapting paths is not sufficient for one to conclude that X should extort Y. If there are other adapting paths where Y ends up at a local optimum, it is unclear what X should do, as discussed above. In the numerical experiment of [17] for the conventional parameters mentioned before, the adapting paths examined do not end up at a local optimum, but formal analysis of the general case is missing.

The same question can be asked for all pcZD strategies, not only the extortionate ones.

Our contribution. We prove the conjecture of [17] in a very strong and general form, and analytically justify the use of extortionate as well as other pcZD strategies against adapting players. We show that in all cases, *all* adapting paths of Y lead to the maximum scores, although the strategy of Y may not end up at the unconditional cooperation. This holds even in some degenerate cases where the analysis of [17] does not apply. Accordingly, as long as Y does not stop at a locally sub-optimal strategy and does not evolve at a speed that goes to 0 as time goes to infinity, the dynamics will always end up at the maximum scores attainable under the linear relation imposed by X. Therefore, it is always "safe" for X to use pcZD strategies, and she will receive her desired score in a very robust way, without knowing which adapting path Y will follow.

As an easy consequence of our main result, if X does not want to take any advantage over Y, but instead is benevolent and wishes to promote mutual cooperation, she is able to do so in all cases, via a "fair" extortionate strategy, where $\chi = 1$, or via a generous strategy. In this way, X enforces the maximum total score of the two players, (R, R), which de facto is equivalent to the unconditional cooperation by both players. This is true even when Y only evolves selfishly and does not care about the total score at all.

Related work. The original setup of Press and Dyson [17] is very different from that of all other studies of ZD strategies so far [1,2,10,11,21,22]. In particular, in [17] there are only two players, one of them uses a fixed ZD strategy and the other changes his strategy over time. While in all other studies, which focus on evolutionary aspects of ZD strategies, there are one or two populations of players and all players can change their strategies over time (the study in [2] has two parts, where the first part considers two players but none changes his strategy, and the second part considers a population of players). Moreover, in [17] only the performances of the ZD strategy and the adaptive player's (non-ZD) strategy against each other matter, while in all other studies the performances of ZD strategies against ZD strategies and the performances of non-ZD strategies against non-ZD strategies are also important.

We follow the original model of [17]. Thus, it is not surprising that our conclusions are consistent with those of [17] and yet seem different from or even opposite to some of the other studies. Indeed, [1] has argued that extortionate strategies are evolutionarily unstable, and [10] has shown via numerical simulations that ZD strategies in general are disfavored by selection when the population is sufficiently large. Generous ZD strategies are first considered by [20] and then studied by [2, 21], which have shown that such strategies are in general evolutionarily stable and disproportionately favored by selection. Thus, among ZD strategies that have been studied, it seems that only generous strategies can be successful from evolutionary aspects, at least when the population is sufficiently large. In contrast, we show that all pcZD strategies, including both extortionate ones and generous ones, are very successful facing an adaptive player in two-player IPD games.

In terms of methodology, most existing studies of ZD strategies either rely on numerical simulations or focus on a finite set of representative strategies; see [1, 10, 17, 21, 22]. A recent work by [11] uses both an analytical framework and numerical simulations to study the evolutionary dynamics of all ZD strategies, when all players apply strategies that may not be ZD but still enforce a linear relation between payoffs of the two players in the repeated game. Our results are derived purely analytically and apply to all pcZD strategies and all adapting paths of the non-ZD player.

Outline of the paper. The results of [17] are recalled in Section 2. In Section 3, we state our main theorem and outline its proof, which is carried out in Section 4. In Section 5, we discuss implications of our results and other problems where our approach can be applied. Appendix A provides some examples on the effect of Y's increased cooperation facing arbitrary strategies of X, and Appendix B explains the key algebraic observations leading to our results.

2 Review of Zero-Determinant Strategies

For computational purposes, it is convenient to measure all scores from the base level P and relative to R. Thus, we first shift down the values of R, T, S, P, s_X, s_Y in [17] by P and then divide them by R (which is the quantity $R-P$ before shifting and is positive), so that the resulting IPD parameters now satisfy

$$P = 0, \quad R = 1, \quad S < 0, \quad T > 1, \quad 0 < S + T < 2. \tag{1}$$

Following [17], we denote an outcome of a one-shot PD game by $xy \in \{CC, CD, DC, DD\}$, with the first letter being the choice of X. As shown in [17], for the study of ZD strategies it is sufficient to consider strategies depending only on the outcome of the last round. Such a strategy for X is described by a tuple $\mathbf{p} = (p_1, p_2, p_3, p_4)$, corresponding to the probabilities that X cooperates in the current move when the previous outcome is CC, CD, DC, and DD, respectively. Symmetrically, a strategy of Y is described by a tuple $\mathbf{q} = (q_1, q_2, q_3, q_4)$, corresponding to the probabilities that Y cooperates when the previous outcome is CC, DC, CD, and DD, respectively. Notice that p_2 and q_2 correspond to different outcomes, representing the different views of X and Y: if we switch the letters of an outcome and let the first letter be the choice of Y, then $\mathbf{q}$ represents the probabilities that Y cooperates when the previous outcome is CC, CD, DC, and DD, respectively.

Each pair of strategies $(\mathbf{p}, \mathbf{q})$ induces a Markov chain with the transition matrix

$$\mathbf{M}(\mathbf{p}, \mathbf{q}) = \begin{pmatrix} p_1q_1 & p_1(1-q_1) & (1-p_1)q_1 & (1-p_1)(1-q_1) \\ p_2q_3 & p_2(1-q_3) & (1-p_2)q_3 & (1-p_2)(1-q_3) \\ p_3q_2 & p_3(1-q_2) & (1-p_3)q_2 & (1-p_3)(1-q_2) \\ p_4q_4 & p_4(1-q_4) & (1-p_4)q_4 & (1-p_4)(1-q_4) \end{pmatrix}, \tag{2}$$

where rows and columns are indexed from the view of X (that is, by the vector (CC, CD, DC, DD)) and $\mathbf{M}(\mathbf{p}, \mathbf{q})_{xy, x'y'}$ represents the probability of seeing outcome $x'y'$ when the previous outcome is xy. The players' scores under $(\mathbf{p},\mathbf{q})$, denoted by $s_X(\mathbf{p},\mathbf{q})$ and $s_Y(\mathbf{p},\mathbf{q})$, are defined to be their expected scores under the stationary distribution $\mathbf{v}(\mathbf{p},\mathbf{q})$, which is treated as a row vector and is multiplied by $\mathbf{M}(\mathbf{p},\mathbf{q})$ on the right (in the degenerate cases, when $\mathbf{M}(\mathbf{p},\mathbf{q})$ has two or more stationary distributions, the scores also depend on the starting outcome).

The fundamental observation of [17] is that the stabilized scores $s_X(\mathbf{p},\mathbf{q})$ and $s_Y(\mathbf{p},\mathbf{q})$, in nearly all cases, are given by

$$s_X(\mathbf{p},\mathbf{q}) = \frac{D(\mathbf{p},\mathbf{q},\mathbf{S}_X)}{D(\mathbf{p},\mathbf{q},\mathbf{1}_4)}, \quad s_Y(\mathbf{p},\mathbf{q}) = \frac{D(\mathbf{p},\mathbf{q},\mathbf{S}_Y)}{D(\mathbf{p},\mathbf{q},\mathbf{1}_4)}, \tag{3}$$

where $\mathbf{S}_X = (1,T,S,0)$, $\mathbf{S}_Y = (1,S,T,0)$, $\mathbf{1}_4 = (1,1,1,1)$, and

$$D(\mathbf{p},\mathbf{q},\mathbf{f}) = \det\begin{pmatrix} -1+p_1q_1 & -1+p_1 & -1+q_1 & f_1 \\ p_3q_2 & p_3 & -1+q_2 & f_2 \\ p_2q_3 & -1+p_2 & q_3 & f_3 \\ p_4q_4 & p_4 & q_4 & f_4 \end{pmatrix} \tag{4}$$

for any $\mathbf{f} = (f_1,f_2,f_3,f_4)$. We have switched the second and third entries of the payoff vectors of [17], as well as the second and third rows of the matrix of [17]. We have also renamed the vector $\mathbf{f}$ so that each f_ℓ still appears in the ℓ-th row. These changes are to simplify the discussion later on (each q_ℓ now appears in the ℓ-th row of the matrix) and have no effect on the left-hand sides of the equations in (3). These equations are valid as long as $D(\mathbf{p},\mathbf{q},\mathbf{1}_4) \neq 0$. We call the strategy $\mathbf{1}_4$ the unconditional cooperation strategy.

For many $(\alpha,\beta,\gamma) \in \mathbb{R}^3$, equations in (3) allow player X to choose a strategy $\mathbf{p}$ so that

$$\alpha s_X + \beta s_Y + \gamma = 0 \tag{5}$$

for any strategy $\mathbf{q}$ of Y. As shown in [17], a suitable strategy of X is given by

$$\mathbf{p} = \big(1+\phi(\alpha+\beta+\gamma), 1+\phi(\alpha S+\beta T+\gamma), \phi(\alpha T+\beta S+\gamma), \phi\gamma\big) \tag{6}$$

for some $\phi \neq 0$ so that $\mathbf{p} \in [0,1]^4$ (if it exists). Thus, a ZD strategy, i.e. $\mathbf{p}$ as above for some α,β,γ and ϕ, satisfies

$$p_2 + p_3 = 1 + 2p_4 - (1-p_1+p_4)(T+S)\,. \tag{7}$$

Whenever $\alpha \neq 0$, the condition (5) is equivalent to the condition

$$s_X - K = \chi(s_Y - K), \tag{8}$$

where $\chi = -\beta/\alpha$, $K = -\gamma/(\alpha+\beta)$ if $\alpha \neq -\beta$, and K does not matter otherwise (if $\alpha = -\beta$ and $\mathbf{p}$ in (6) lies in $[0,1]^4$ for some $\phi \neq 0$, then $\gamma = 0$ in (5)). A suitable strategy of X is then given by

$$\mathbf{p} = \big(1-\phi(\chi-1)(1-K), 1-\phi(\chi T - S - (\chi-1)K), \phi(T-\chi S+(\chi-1)K), \phi(\chi-1)K\big)\,, \tag{9}$$

where ϕ has been scaled up by α from equation (6). The $K=0$ case of (9) is equation (12) of [17] with typos corrected (ϕ in (12) of [17] should be replaced by $(P\text{–}S)\phi$ to be consistent with (11); this typo carries over to (13)). The scores of X and Y are positively correlated if $\chi \geq 1$ (the cases $\chi \in (0,1)$ do not correspond to any ZD strategy; see Section 4). An extortionate strategy in the terminology of [17] is a strategy $\mathbf{p}$ of X enforcing equation (8) with $K=0$ and some fixed $\chi \geq 1$; it ensures that X gets a higher share of the total payoff (above the base level $P=0$). A generous strategy in the terminology of [2] and [21] is a strategy $\mathbf{p}$ of X enforcing (8) with $K=1$ and some fixed $\chi \geq 1$; it ensures that X takes a higher share of the total *loss* below the total unconditional-cooperation reward $2R = 2$. The value of χ is called the extortion factor in the first case and the generosity factor in the second case. For $\chi = 1$ (in which case K is irrelevant), the strategy $\mathbf{p}$ is called fair.

By (3), the players' payoffs when X plays a ZD strategy (9) and Y unconditionally cooperates are given by

$$s_X^*(\chi, K) = 1 + (\chi-1)\frac{(1-K)(T-1)}{T-1+\chi(1-S)}, \quad s_Y^*(\chi, K) = 1 - (\chi-1)\frac{(1-K)(1-S)}{T-1+\chi(1-S)} \tag{10}$$

and in particular are independent of the value ϕ in (9). Since $s_X^*(\chi, K)+s_Y^*(\chi, K)$ cannot exceed $2R = 2$, there are no values of ϕ so that $\mathbf{p} \in [0,1]^4$ (i.e. the relation (8) cannot be enforced) if $(\chi-1)(1-K) < 0$. Since $s_X^*(\chi, K)+s_Y^*(\chi, K) = 2$ if $\chi = 1$ or $K = 1$, the maximal possible scores for a fixed fair or generous ZD strategy of X are the unconditional cooperation scores (10). For $K = 0$, (10) provides the maximal possible scores for an extortionate ZD strategy, as stated in equation (14) of [17]. As a corollary of our analysis of adapting paths, we show that s_X and s_Y are maximized for any fixed positively correlated ZD (pcZD) strategy $\mathbf{p}$ of X whenever Y is unconditionally cooperative.

Even a priori knowledge that $s_Y(\mathbf{q})$ is maximized at the unconditionally cooperative strategy $\mathbf{q} = \mathbf{1}_4$ does not imply that there are adapting paths for Y that eventually lead to this score, as the function $s_Y(\mathbf{q})$ could have local peaks. For a fixed extortionate strategy, [17] conjectures that

(a) there exist in all cases adapting paths for Y along which the directional derivatives of s_Y with respect to q_1, q_2, q_3, q_4 are always positive, and thus

(b) Y would evolve to a strategy that achieves the maximum possible scores.

Numerical evidence for these conjectures, with the conventional values of R, T, S, P and the unconditionally non-cooperative initial strategy $(0, 0, 0, 0)$ of Y, is provided in [17]. We not only confirm these conjectures, but extend them to the most general form for all pcZD strategies. We prove that

(**a**) for all pcZD strategies of X in all IPD games (with the standard restrictions on the parameters given by (11)), the directional derivatives of s_Y with respect to q_1, q_2, q_3, q_4 are everywhere positive (with rare exceptions when some, but not all of them, are zero), and

(**b**) for all pcZD strategies in all IPD games, all adapting paths of Y lead to the maximum possible scores for both players given by (10), but not necessarily to the unconditionally cooperative strategy of Y, in a finite time, even though some adapting paths may pass through degenerate points where the score $s_Y(\mathbf{q})$ is not continuous in $\mathbf{q}$.

In particular, the fair ($\chi = 1$) and the generous ($K = 1$) ZD strategies against an adapting player always lead to the optimal outcome (the total score $s_X + s_Y$ at the unconditional cooperation level of $2R = 2$).

3 Main Result

We first define adapting paths for Y.

Definition. An adapting path for Y is a smooth map $\lambda\colon [0, \tau] \longrightarrow [0,1]^4$, for some $\tau \in \mathbb{R}$, such that

(A1) $s_Y(\lambda(t_1)) < s_Y(\lambda(t_2))$ whenever $0 \le t_1 < t_2 \le \tau$;

(A2) there is no smooth map $\tilde{\lambda} : [0, \tilde{\tau}] \longrightarrow [0,1]^4$ such that $\tilde{\lambda}$ satisfies the restriction (A1) and $\lambda([0,\tau]) = \tilde{\lambda}([0, \tilde{\tau}'])$ for some $\tilde{\tau}' < \tilde{\tau}$.

In this definition, the input of λ is time and the output is a mixed strategy of Y. The restriction (A1) means that Y improves his own utility by changing his strategies; the restriction (A2) means that Y does not stop at a locally sub-optimal strategy. The fact that τ is finite reflects the requirement that Y does not adapt under a speed that goes to 0 as time goes to infinity. Below we state our main theorem.

Theorem. *Let* $\mathbf{p}$ *be a pcZD strategy of player* X *in an IPD game with the payoffs satisfying*

$$S < P < R < T, \qquad 2P < S + T < 2R. \tag{11}$$

Every adapting path for the strategy of player Y *leads to a strategy* $\mathbf{q}$ *with* $(q_1, q_2) = (1, 1)$ *if* $p_1 < 1$ *and with* $q_1 = 1$ *if* $p_1 = 1$*. This strategy for* Y *is de facto equivalent to the unconditional cooperation strategy and maximizes the stationary payoff scores* s_X *and* s_Y *among all the possible strategies of* Y *(for the given strategy* $\mathbf{p}$ *of* X*); these scores are given by equation (10). Furthermore, an adapting path for* Y *always exists.*

Outline of the proof. In order to understand the adapting paths of Y, we need to characterize the partial derivatives of s_Y with respect to $\mathbf{q}$. Given a ZD strategy $\mathbf{p}$, s_X and s_Y are linearly correlated, either positively or negatively, where the latter corresponds to $\chi < 0$ in the definition of ZD strategies. Thus, equivalently we can characterize the partial derivatives of s_X. The key to our result is an observation that a positive multiple of the partial derivative of s_X with respect to each q_i splits into factors that are independent of q_i and are linear in each of the other q_j's, no matter what $\mathbf{p}$ is (the mathematical deduction of the factorization is provided in Appendix B and the result is used in equation (14)). Given the factorization, the q_i-th partial of s_X on the entire 4-cube of the possible strategies $\mathbf{q}$ is nonnegative if the factors are all nonnegative at the 8 corners of the 3-cube of the other three variables q_j with $j \neq i$. These partials are often positive (or at least nonnegative), reflecting the fact that increased cooperation by Y generally helps X, but this is not always the case for an arbitrary non-ZD strategy $\mathbf{p}$ of X; see Appendix A. For a typical ZD strategy, depending on whether s_X and s_Y are positively or negatively correlated, the q_i-partials of s_X specify whether it is desirable for Y to increase his cooperation. For all pcZD strategies in all IPD games (with standard restrictions on the parameters as specified by equation (11)), by examining the factorization of the partials one can see that it is indeed desirable for Y to increase his cooperation. In these cases, an adapting player Y moves toward the unconditionally cooperative strategy or an effectively equivalent one. On the other hand, against negatively correlated ZD strategies $\mathbf{p}$ of X, Y would generally reduce his cooperation, helping his score and hurting X's.

4 Proof of the Main Theorem

We first note some necessary conditions on (χ, K) for a pcZD strategy (9) to exist and the resulting properties of $\mathbf{p}$. As $\phi = 0$ in (9) would correspond to taking $\alpha, \beta, \gamma = 0$ in (5), which is meaningless, we have $\phi \neq 0$ in (9). As noted after (10), $(\chi - 1)(1 - K) \geq 0$; since $p_1 \leq 1$, this implies that either $\phi > 0$, or $\chi = 1$, or $K = 1$. Since $p_3 \geq 0$ and (S, T) satisfies (1), either of the last two assumptions also implies that $\phi > 0$. So $\phi > 0$ always. Since $p_4 \geq 0$, it follows that $(\chi - 1)K \geq 0$ and so $\chi \geq 1$ and $K \in [0, 1]$ if $\chi > 1$ (if $\chi = 1$, K does not matter). Combining these observations with the conditions in (1), we conclude that

$$\chi \geq 1, \quad 0 \leq K \leq 1, \quad p_1 > p_2, \quad p_3 > p_4\,.$$

Thus,

$$p_1, \hat{p}_2, p_3, \hat{p}_4 > 0, \tag{12}$$

(q_1, q_2) \ (q_3, q_4)	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,0)$	$\hat{p}_2 + p_4$	$p_1 p_4 + 2\hat{p}_2 + p_3 \hat{p}_4$	$\hat{p}_1 p_2 + \hat{p}_2 \hat{p}_3 + 2p_4$	$2(1-(p_1-p_3)(p_2-p_4))$
$(0,1)$	$(\hat{p}_2 + p_4)p_3$	$p_1 p_3 + \hat{p}_2(\hat{p}_4 + 2p_3)$	$\hat{p}_1 p_3 + (\hat{p}_2 + 2p_3)p_4$	$p_1 \hat{p}_2 + \hat{p}_1 \hat{p}_4 + 2p_3$
$(1,0)$	$\hat{p}_1(\hat{p}_2 + p_4)$	$\hat{p}_1(2\hat{p}_2 + p_3) + \hat{p}_2 p_4$	$\hat{p}_1(\hat{p}_3 + 2p_4) + p_2 p_4$	$2\hat{p}_1 + p_2 p_3 + \hat{p}_3 p_4$
$(1,1)$	0	$(\hat{p}_1 + p_3)\hat{p}_2$	$(\hat{p}_1 + p_3)p_4$	$\hat{p}_1 + p_3$

Table 1: The values of the determinant $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4)$ at the corners of $[0,1]^4$; $\hat{p}_i \equiv 1 - p_i$.

(q_1, q_2) \ (q_3, q_4)	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,0)$	none	none	$\mathbf{p} = (1,0,1,0)$	none
$(0,1)$	none	none	$(p_1, p_4) = (1,0)$	none
$(1,0)$	$p_1 = 1$	$(p_1, p_4) = (1,0)$	$(p_1, p_4) = (1,0)$	$\mathbf{p} = (1,0,1,0)$
$(1,1)$	all	none	$p_4 = 0$	none

Table 2: The pcZD strategies $\mathbf{p}$ for which $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) = 0$ for the specified values of $\mathbf{q} \in \{0,1\}^4$.

where $\hat{p}_i = 1 - p_i$.

We next derive some properties of the determinants in (3). An important feature of the function $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4)$ defined by (4) is that it is linear in each variable q_1, q_2, q_3, q_4 separately. Thus, in order to check whether $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) = 0$ for any of the possible values of $\mathbf{q}$, it is sufficient to consider the values of $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4)$ only for the extremal values of q, that is, for $q_1, q_2, q_3, q_4 = 0, 1$. This is simple to do; the results are summarized in Table 1. We can see that for every strategy $\mathbf{p}$ of X, $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) \geq 0$ at the 16 corners of $[0,1]^4$, and thus $D(\mathbf{p}, \mathbf{q}, \mathbf{1}) \geq 0$ everywhere on $[0,1]^4$. For a generic strategy $\mathbf{p}$, $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) = 0$ at only one of the 16 corners, $\mathbf{q} = (1,1,0,0)$, which implies that $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) > 0$ everywhere else on $[0,1]^4$.

By Table 1, $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) = 0$ at all 16 corners only for the strategy $\mathbf{p} = (1,1,0,0)$. By (12), this is not a pcZD strategy. Thus, for every pcZD strategy $\mathbf{p}$, $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) > 0$ on $(0,1)^4$. Indeed, if $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) = 0$ for any interior point $\mathbf{q} \in (0,1)^4$, it must be 0 at all 16 corners, due to the linearity of $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4)$ in each q_i. Table 2 characterizes, at each of the 16 corners, the pcZD strategies for which $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) = 0$. Thus, it completely determines the vanishing locus of $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) = 0$ for every pcZD strategy $\mathbf{p}$; the results are summarized in the following proposition.

Proposition 1. *For every* $\mathbf{p} \in [0,1]^4$, $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) \geq 0$ *for all* $\mathbf{q} \in [0,1]^4$. *For a pcZD strategy* $\mathbf{p}$ *and a strategy* $\mathbf{q}$, $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4) = 0$ *if and only if*

(D1) $\mathbf{q} = (1,1,0,0)$, *or*

(D2) $p_4 = 0$ *and* $(q_1, q_2, q_4) = (1,1,0)$, *or*

(D3) $p_1 = 1$ *and* $(q_1, q_3, q_4) = (1,0,0)$, *or*

(D4) $(p_1, p_4) = (1,0)$ *and either* $(q_1, q_4) = (1,0)$, *or* $(q_1, q_2, q_3) = (1,0,0)$, *or* $(q_2, q_3, q_4) = (1,1,0)$, *or*

(D5) $\mathbf{p} = (1,0,1,0)$ *and either* $(q_1, q_2) = (1,0)$ *or* $(q_3, q_4) = (1,0)$.

We now derive some properties of the partial derivatives, which are well defined as long as

$\mathbf{q}_{-\ell}$	$-\langle M_1\rangle$	$\langle M_2\rangle$	$-\langle M_3\rangle$	$\langle M_4\rangle$
$(0,0,0)$	0	0	p_4	$\hat{p}_2$
$(0,0,1)$	$\hat{p}_2 p_4$	$\hat{p}_2\hat{p}_4$	$p_1p_4+p_3\hat{p}_4$	$\hat{p}_1p_2+\hat{p}_2\hat{p}_3$
$(0,1,0)$	p_2p_4	$\hat{p}_2p_4$	p_3p_4	$\hat{p}_2p_3$
$(0,1,1)$	$p_2p_3+\hat{p}_3p_4$	$p_1\hat{p}_2+\hat{p}_1\hat{p}_4$	p_1p_3	$\hat{p}_1p_3$
$(1,0,0)$	0	0	$\hat{p}_1p_4$	$\hat{p}_1\hat{p}_2$
$(1,0,1)$	$\hat{p}_2p_3$	$\hat{p}_1\hat{p}_2$	$\hat{p}_1p_3$	$\hat{p}_1\hat{p}_3$
$(1,1,0)$	p_3p_4	$\hat{p}_1p_4$	0	0
$(1,1,1)$	p_3	$\hat{p}_1$	0	0

Table 3: The values of the determinants $\langle M_\ell\rangle$ at the corners of $[0,1]^3$ for $\mathbf{q}_{-\ell}$; $\hat{p}_i\equiv 1-p_i$.

$D(\mathbf{p},\mathbf{q},\mathbf{1})\neq 0$. By (3) and the Quotient Rule,

$$D(\mathbf{p},\mathbf{q},\mathbf{1}_4)^2\cdot\frac{\partial s_X}{\partial q_\ell}=D(\mathbf{p},\mathbf{q},\mathbf{1}_4)\cdot\frac{\partial D(\mathbf{p},\mathbf{q},\mathbf{S}_X)}{\partial q_\ell}-\frac{\partial D(\mathbf{p},\mathbf{q},\mathbf{1}_4)}{\partial q_\ell}\cdot D(\mathbf{p},\mathbf{q},\mathbf{S}_X). \tag{13}$$

Since the functions $\mathbf{q}\longrightarrow D(\mathbf{p},\mathbf{q},\mathbf{1}_4)$ and $\mathbf{q}\longrightarrow D(\mathbf{p},\mathbf{q},S_X)$ are linear in each q_ℓ, the right-hand side of (13) is independent of q_ℓ and is of degree at most 2 in each of the other three variables q_j with $j\neq\ell$. It turns out that the right-hand side of (13) splits into two factors, each of which is linear in each q_j with $j\neq\ell$; we give a conceptual reason for this and describe the two factors explicitly in Appendix B. Thus, the sign of the right-hand side of (13) on $[0,1]^4$ is completely determined by the signs of each of the two factors at the 8 corners of $[0,1]^3$ for the three variables q_j with $j\neq\ell$, provided these signs are the same for each of the two factors.

If $\mathbf{p}$ satisfies (7), in particular if $\mathbf{p}$ is a pcZD strategy, the second factor described in (22) of Appendix B splits further and

$$D(\mathbf{p},\mathbf{q},\mathbf{1}_4)^2\cdot\frac{\partial s_X}{\partial q_\ell}=\big(\hat{p}_2-\hat{p}_1S+p_4(1-S)\big)\langle M_\ell\rangle\mathfrak{d}_\ell, \tag{14}$$

where $\langle M_\ell\rangle$ is the determinant of the $(\ell,4)$-minor of the matrix in (4) and $\mathfrak{d}_\ell$ is linear in each q_j with $j\neq\ell$ (as is $\langle M_\ell\rangle$); this statement can be verified directly (using *Mathematica*, for example). By (12) and the fact that $S<0$,

$$\hat{p}_2-\hat{p}_1S+p_4(1-S)>0$$

for every pcZD strategy $\mathbf{p}$. Thus, the sign behavior of each q_ℓ-partial of s_X is determined by the signs of $\langle M_\ell\rangle$ and $\mathfrak{d}_\ell$ at the 8 corners of $[0,1]^3$ for the q_j's with $j\neq\ell$.

The eight signs of the two factors are described by Tables 3 and 5. Since $(-1)^\ell\langle M_\ell\rangle,(-1)^\ell\mathfrak{d}_\ell\geq 0$ at all 8 corners, $\langle M_\ell\rangle\mathfrak{d}_\ell\geq 0$ on $[0,1]^3$. Tables 4 and 6 characterize, at each of the 8 corners, the pcZD strategies for which $\langle M_\ell\rangle,\mathfrak{d}_\ell=0$, and thus completely determine the vanishing locus of (14). Since s_X and s_Y are positively correlated, the results apply equally well to the partials of s_Y and are summarized in the following proposition (we started by describing the partials of s_X as they are more likely to be positive for general, not necessarily pcZD, strategies).

Proposition 2. *For a pcZD strategy* $\mathbf{p}$,

$$\frac{\partial s_Y}{\partial q_1},\frac{\partial s_Y}{\partial q_2},\frac{\partial s_Y}{\partial q_3},\frac{\partial s_Y}{\partial q_4}\geq 0 \tag{15}$$

for every $\mathbf{q}\in[0,1]^4$ *such that* $D(\mathbf{p},\mathbf{q},\mathbf{1}_4)\neq 0$. *Moreover,*

$\mathbf{q}_{-\ell}$	$-\langle M_1\rangle$	$\langle M_2\rangle$	$-\langle M_3\rangle$	$\langle M_4\rangle$
$(0,0,0)$	all	all	$p_4=0$	none
$(0,0,1)$	$p_4=0$	none	none	$(p_1,p_3)=(1,1)$ or $(p_2,p_3)=(0,1)$
$(0,1,0)$	$p_2=0$ or $p_4=0$	$p_4=0$	$p_4=0$	none
$(0,1,1)$	$(p_2,p_3)=(0,1)$ or $(p_2,p_4)=(0,0)$	none	none	$p_1=1$
$(1,0,0)$	all	all	$p_1=1$ or $p_4=0$	$p_1=1$
$(1,0,1)$	none	$p_1=1$	$p_1=1$	$p_1=1$ or $p_3=1$
$(1,1,0)$	$p_4=0$	$p_1=1$ or $p_4=0$	all	all
$(1,1,1)$	none	$p_1=1$	all	all

Table 4: The pcZD strategies $\mathbf{p}$ for which $\langle M_\ell\rangle=0$ for the specified values of $\mathbf{q}_{-\ell}\in\{0,1\}^3$.

$\mathbf{q}_{-\ell}$	$-\mathfrak{d}_1$	$\mathfrak{d}_2$	$-\mathfrak{d}_3$	$\mathfrak{d}_4$
$(0,0,0)$	$p_1+\theta\hat{p}_1$	$p_3+\theta\hat{p}_3$	$p_2+\theta\hat{p}_2$	$p_4+\theta\hat{p}_4$
$(0,0,1)$	$(2-\theta)p_1+\hat{p}_3+(p_1-p_2)$	$p_3+\theta\hat{p}_3+(2-\theta)(p_3-p_4)$	$(2-\theta)p_2+\hat{p}_3+\theta(p_1-p_2)$	$p_2+\theta\hat{p}_4$
$(0,1,0)$	$\theta\hat{p}_1+p_2$	$p_2+\theta\hat{p}_3$	$\theta\hat{p}_2+p_3$	$p_3+\theta\hat{p}_4$
$(0,1,1)$	$(2-\theta)p_1+\hat{p}_3$	$p_2+\theta\hat{p}_3+(2-\theta)(p_3-p_4)$	$1-(\theta-1)p_2+\theta(p_1-p_2)$	$p_2+\theta\hat{p}_3+(1+\theta)(p_3-p_4)$
$(1,0,0)$	$\hat{p}_2+(2-\theta)p_4$	$\theta\hat{p}_1+p_3$	$\theta\hat{p}_1+p_2$	$\theta\hat{p}_1+p_4$
$(1,0,1)$	$(2-\theta)p_1+\hat{p}_2$	$\hat{p}_2+(2-\theta)p_3$	$(2-\theta)p_2+\hat{p}_3$	$\theta\hat{p}_1+p_2$
$(1,1,0)$	$\hat{p}_1+(2-\theta)p_4$	$\hat{p}_3+(2-\theta)p_4$	$\hat{p}_2+(2-\theta)p_4$	$\hat{p}_2+(2-\theta)p_4$
$(1,1,1)$	$1-(\theta-1)p_1$	$1-(\theta-1)p_3$	$1-(\theta-1)p_2$	$1-(\theta-1)p_4$

Table 5: The values of the factors $\mathfrak{d}_\ell$ at the corners of $[0,1]^3$ for $\mathbf{q}_{-\ell}$; $\hat{p}_i\equiv 1-p_i$, $\theta\equiv T+S$.

- $\frac{\partial s_Y}{\partial q_1}=0$ *if and only if* $(q_3,q_4)=(0,0)$, *or* $(p_4,q_4)=(0,0)$, *or* $(p_4,q_2,q_3)=(0,0,0)$, *or* $(p_2,q_2,q_4)=(0,0,0)$, *or* $(p_2,p_4,q_2)=(0,0,0)$, *or* $(p_2,p_3,q_2,q_3)=(0,1,0,1)$*;*
- $\frac{\partial s_Y}{\partial q_2}=0$ *if and only if* $(q_3,q_4)=(0,0)$, *or* $(p_4,q_4)=(0,0)$, *or* $(p_1,q_1)=(1,1)$, *or* $(p_2,p_3,q_1,q_3,q_4)=(0,1,0,1,0)$*;*
- $\frac{\partial s_Y}{\partial q_3}=0$ *if and only if* $(q_1,q_2)=(1,1)$, *or* $(p_4,q_4)=(0,0)$, *or* $(p_1,q_1)=(1,1)$, *or* $(p_2,p_3,q_1,q_2,q_4)=(0,1,1,0,1)$*;*
- $\frac{\partial s_Y}{\partial q_4}=0$ *if and only if* $(q_1,q_2)=(1,1)$, *or* $(p_1,q_1)=(1,1)$, *or* $(p_1,q_2,q_3)=(1,1,1)$, *or* $(p_3,q_1,q_3)=(1,1,1)$, *or* $(p_1,p_3,q_3)=(1,1,1)$, *or* $(p_2,p_3,q_2,q_3)=(0,1,0,1)$.

Corollary 1. *For any pcZD strategy* $\mathbf{p}$ *and any strategy* $\mathbf{q}$ *such that* $D(\mathbf{p},\mathbf{q},\mathbf{1}_4)\neq 0$, $\frac{\partial s_Y}{\partial q_j}>0$ *for some* $j=1,2,3,4$.

We can now derive some properties of Y's adapting paths. By equation (15) and Corollary 1, the value of at least one q_i increases at each point of an adapting path, at least outside of the vanishing locus of the function $\mathbf{q}\longrightarrow D(\mathbf{p},\mathbf{q},\mathbf{1}_4)$ described by Proposition 1 (could be different q_i at different points of the same adapting path). A smooth map $\lambda\colon[0,\tau]\longrightarrow[0,1]^4$ satisfying (A1) and $\lambda(\tau)=\mathbf{q}'$ can be extended further and thus cannot be an adapting path, unless

$$D(\mathbf{p},\mathbf{q}',\mathbf{1}_4)=0 \quad \text{or} \quad \frac{\partial s_Y}{\partial q_i}\Big|_{\mathbf{q}'}=0 \quad \forall\, i \text{ s.t. } q_i'<1, \tag{16}$$

$\mathbf{q}_{-\ell}$	$-\mathfrak{d}_1$	$\mathfrak{d}_2$	$-\mathfrak{d}_3$	$\mathfrak{d}_4$
$(0,0,0)$	none	none	none	none
$(0,0,1)$	none	none	none	none
$(0,1,0)$	$(p_1,p_2)=(1,0)$	$(p_2,p_3)=(0,1)$	none	none
$(0,1,1)$	none	none	none	none
$(1,0,0)$	none	none	$(p_1,p_2)=(1,0)$	$(p_1,p_4)=(1,0)$
$(1,0,1)$	none	none	$(p_2,p_3)=(0,1)$	$(p_1,p_2)=(1,0)$
$(1,1,0)$	$(p_1,p_4)=(1,0)$	$(p_3,p_4)=(1,0)$	none	none
$(1,1,1)$	none	none	none	none

Table 6: The pcZD strategies $\mathbf{p}$ for which $\mathfrak{d}_\ell=0$ for the specified values of $\mathbf{q}_{-\ell}\in\{0,1\}^3$.

and

$$\frac{\partial s_Y}{\partial q_i}\Big|_{\mathbf{1}_4}=0 \quad \forall\, i \text{ s.t. } q_i'<1. \tag{17}$$

The conditions (16) describe the singular points and the points with all relevant partial derivatives vanishing. Such points are still not local optima if (17) is not satisfied, because of the linearity of the factors of these partial derivatives as we previously discussed. Only if $\mathbf{q}'$ satisfies (16) *and* (17), it may not be possible to extend λ further. If $D(\mathbf{p},\mathbf{q}',\mathbf{1}_4)\neq 0$ and $\mathbf{q}'$ satisfies the second condition in (16) and the condition in (17), then λ cannot be further extended. However, if $D(\mathbf{p},\mathbf{q}',\mathbf{1}_4)=0$ and $\mathbf{q}'$ satisfies (17), λ may still be extendable.

We now describe when $\mathbf{q}'$ satisfies (16) and (17). The set of $\mathbf{q}'$ satisfying the first condition in (16) is provided by Proposition 1. The second condition in (16) and the one in (17) can be studied either from Proposition 2 or directly from Tables 4 and 6. For example, since $\frac{\partial s_Y}{\partial q_1}\neq 0$ at $\mathbf{q}=\mathbf{1}_4$, no $\mathbf{q}'$ with $q_1'<1$ satisfies both (16) and (17). Considering the 7 possibilities for the nonempty set $\{j\neq 1\colon\, q_j'<1\}$ with $q_1'=1$, we find that $\mathbf{q}'$ satisfies (16) and (17) if and only if

(T1) $(q_1',q_2')=(1,1)$ or

(T2) $p_1=1$ and $q_1'=1$.

The situations corresponding to (T1) and (T2) are depicted respectively in the two diagrams in Figure 2, which indicate the possible flows of the game: under each condition and at any one of the four outcomes, the arrows indicates where the game can go. The solid arrows show the flows that always exist with nonzero probability. At least one of the two dashed flows leaving from the same vertex exists as well. However, the one that is more to the right can be reduced by increasing the value of q_i corresponding to the given vertex; doing so increases both s_X and s_Y and thus will eventually be carried out by player Y, if the stationary distribution passes through that vertex. The dotted flows need not exist, depending on $\mathbf{p}$.

For most values of the undetermined parameters q_i ($i=3,4$ in the first diagram and $i=2,3,4$ in the second), there is a unique stationary distribution (CC/DC combination in the first diagram and CC in the second); since $D(\mathbf{p},\mathbf{q}',\mathbf{1}_4)\neq 0$ in these cases, λ cannot be extended. In these situations, the terminal strategy of Y is equivalent to the unconditional cooperation strategy, since the undetermined parameters q_i have no effect on the outcome. For other values of the undetermined parameters q_i, $D(\mathbf{p},\mathbf{q}',\mathbf{1}_4)=0$ and there may be one or two (if $p_1=1$) stationary distributions in addition to the above one. If the game starts in the above stationary distribution, it will stay there forever. If it starts in a different stationary distribution, increasing q_4 from 0 in the first case and q_2 and/or q_4 in the second case would move the game into the main stationary distribution. This

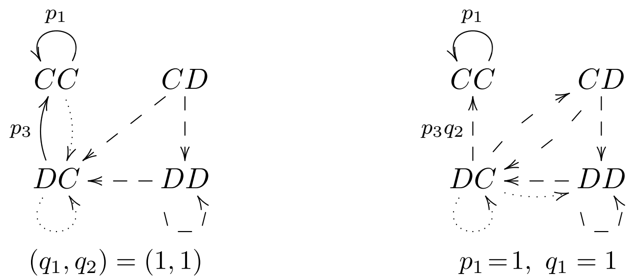

Figure 2: Possible transitions of IPD for the final strategies for adapting player Y; dashed arrows from DD to CC are omitted.

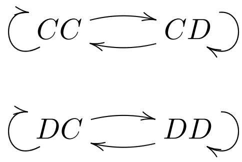

Figure 3: Possible transitions of IPD in the $\phi = 0$ case of (9).

would increase the score of X (as evident from the two diagrams) and thus the score of Y (as they are positively correlated). With either starting situation, the terminal strategy of Y will again be de facto equivalent to the unconditional cooperation strategy. Thus, equation (10) describes the scores corresponding to these situations. Since all adapting paths lead to these scores, they are the maximal possible. Therefore, our theorem holds.

Remark: a special non-ZD strategy. We showed at the beginning of Section 4 that $\phi \neq 0$ in (9). However, the case $\phi = 0$, i.e. $\mathbf{p} = (1,1,0,0)$, is formally allowed by [17]. For this strategy of X and most strategies $\mathbf{q}$ of Y, there are two stationary distributions,

$$\mathbf{v}_C = \frac{(q_3, 1-q_1, 0, 0)}{1 - q_1 + q_3}, \qquad \mathbf{v}_D = \frac{(0, 0, q_4, 1-q_2)}{1 - q_2 + q_4};$$

they are depicted in Figure 3. Under either distribution, Y benefits from evolving to the fully non-cooperative strategy. The same holds for degenerate values of $\mathbf{q}$ as well. Thus, this strategy $\mathbf{p}$ does not work out well for X. This is not too surprising, as $\mathbf{p}$ is not a ZD strategy. As noted at the beginning of Section 4, when $\phi = 0$, equation (11) in [17] imposes no condition on the scores, and the analysis of [17], including equation (14) in [17], does not apply.

5 Discussion

We have formally shown that the extortionate, the generous, and other pcZD strategies are extremely robust at achieving high scores for a sentient player against an adapting opponent. In particular, a sentient player X can force an adapting player Y into effectively unconditional cooperation in all cases: Y's strategy may not evolve to $(1,1,1,1)$, but in the long run Y will always cooperate, because the outcomes at which Y may not cooperate (i.e. $q_i < 1$) never occur when Y stops evolving and the Markov stationary distribution is established. So, the score of X ends up being the same as if Y were unconditionally cooperative and is the maximum possible score for a given pcZD strategy $\mathbf{p}$ of X.

Furthermore, it is immediate from equation (10) that the score for the first player increases with χ if $K < 1$, while the total score $s_X + s_Y$ decreases with χ. The latter is maximized at $\chi = 1$ or

$K = 1$, with each player receiving the mutual cooperation reward $R = 1$. In particular, the $\chi = 1$ and $K = 1$ outcomes are the most desirable from the point of view of social welfare or that of a generous player, but least desirable from the point of view of an extortionate player: there is always some social welfare "burnt" when one player tries to extort the other.

Finally, the approach of this paper can be used to analyze situations when player Y chooses to maximize the relative payoff s_Y/s_X, instead of s_Y. Since

$$\frac{s_Y}{s_X} = \frac{D(\mathbf{p}, \mathbf{q}, S_Y)}{D(\mathbf{p}, \mathbf{q}, S_X)}$$

according to (3), Appendix B shows that all the relevant derivatives are still products of factors that are linear in each of the q_j's. However, this is a question of a different flavor from those considered in this paper (which focus on the effect of pcZD strategies against a payoff-maximizing adapting player), and we do not pursue it here.

Acknowledgments

The authors would like to thank Freeman J. Dyson for explaining [17] and the referees for the comments and suggestions that helped generalize earlier results and streamline the argument. Both authors are also grateful to the IAS School of Mathematics for its hospitality, which lead to part of this work. The first author was partially supported by the Zurich Financial Services and NSF grant CCF-0832797; the second author was partially supported by the IAS Fund for Math.

A Some Examples on the Effect of Increased Cooperation

Although intuitively the more Y cooperates the better X's payoff is, for arbitrary strategies of X this may not be true. We now give examples of strategies $\mathbf{p}$ and $\mathbf{q}$ for X and Y, respectively, so that increasing one of the cooperation variables q_ℓ of Y *hurts* the stabilized score s_X, $s_X + s_Y$, or even both s_X and s_Y. The first example is fairly simple; it is intended to indicate why these phenomena can happen. By Proposition 2, they cannot happen for pcZD strategies.

The basic reason why Y's increasing cooperation in response to some outcomes may hurt X's long-term score is that $T > R = 1$, i.e. the outcome DC is better for X than CC. In the first diagram in Figure 4, if Y were to reduce q_1 from 1, as shown by the dotted arrow in the second diagram, the game would move to CD and then to DC, so X gets hurt in the first step, but gains after that. This situation has two stationary distributions and so is degenerate. The degeneracy is removed by making p_3 a small positive number instead of 0, as shown by the dashed arrow in the second diagram. With $(T, S) = (2, -1/2)$, i.e. the conventional IPD values in our notation,

$$\frac{\partial}{\partial q_1} s_X(q_1, 1, 1, 1)\big|_{q_1=1} = -8.5\,;$$

in particular, the maximum of s_X is not at $\mathbf{q} = (1, 1, 1, 1)$.

The players' total utility may also be reduced if Y increases cooperation. The basic reason is that $T + S < 2R = 2$, i.e. CC is better for $X + Y$ than DC or CD. In the first diagram in Figure 5, with the game flows shown by the solid arrows, if Y were to reduce q_2 from 1 (the dotted arrow), the game would move from DC to DD and then to CC (so $X + Y$ get hurt in the first step, but gain after that). This situation has two stationary distributions and so is degenerate. The degeneracy is removed by making $1 - p_1$ small (the dashed arrow). With $(T, S) = (2, -1/2)$,

$$\frac{\partial}{\partial q_2} s_X(1, q_2, 1, 1)\big|_{q_2=1} + \frac{\partial}{\partial q_2} s_Y(1, q_2, 1, 1)\big|_{q_2=1} = -3.5;$$

Figure 4: In the second diagram, $\mathbf{p} = (1, 0, .1, 0)$, $\mathbf{q} = (q_1, 1, 1, 1)$.

Figure 5: In the first diagram, $\mathbf{p}=(.9, 1, 0, 1)$; in the second diagram, $\mathbf{p}=(.9, 0, 1, 1)$.

in particular, the maximum of s_X+s_Y is not at $\mathbf{q}=(1, 1, 1, 1)$.

Furthermore, the utility of *each* player may be reduced if Y increases cooperation. The basic reason is that $R > (T+S)/2$, i.e. CC is better for X and Y than DC and CD half time each. In the second diagram above, given the solid arrows, if Y were to reduce q_3 from 1 (dotted arrow), the game would move from DC/CD cycle to DD and then to CC (so X and Y get hurt in the first step, but gain after that). This situation is degenerate. To remove the degeneracy, make $1-p_1$ small (dashed arrow). With $(R, T, S)=(2, 4, -1/2)$,

$$\frac{d}{dq_3} s_X(1, 0, q_3, 1)\Big|_{q_3=1} = \frac{d}{dq_3} s_Y(1, 0, q_3, 1)\Big|_{q_3=1} = -0.875.$$

B Some Algebraic Observations

We now give a conceptual reason behind our approach, i.e. that the partial derivatives of s_X and s_Y over $\mathbf{q}$ multiplied by $D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4)^2$ split into factors each of which is linear in each of the variables q_i (and each p_i as well by symmetry).

For any $n\times n$ real matrix M, let $\langle M\rangle$ denote the determinant of M. If in addition $k \in \{1, \dots, n\}$, the k values $i_1, \dots, i_k \in \{1, \dots, n\}$ are all distinct, and $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$, let

$$\langle M; \mathbf{x}_1, \dots, \mathbf{x}_k\rangle_{i_1,\dots,i_k} \in \mathbb{R}$$

denote the determinant of the matrix obtained from M by replacing its columns numbered $i_1, \dots, i_k$ by the column vectors $\mathbf{x}_1, \dots, \mathbf{x}_k$. We note that

$$\det\Big(\big(\langle M; \mathbf{x}_j\rangle_{i_l}\big)_{j,l=1,\dots,k}\Big) = \langle M\rangle^{k-1}\langle M; \mathbf{x}_1, \dots, \mathbf{x}_k\rangle_{i_1,\dots,i_k} \tag{18}$$

for the following reason. Since the subspace of nondegenerate (invertible) matrices is dense and both sides of (18) are continuous in M, it is sufficient to verify (18) under the assumption that the columns of M span $\mathbb{R}^n$. Since both sides are linear and anti-symmetric in the inputs $\mathbf{x}_1, \dots, \mathbf{x}_k$ and vanish if some $\mathbf{x}_j$ equals to a column of M not numbered $i_1, \dots, i_k$, it is sufficient to verify (18) only for the case with $\mathbf{x}_1, \dots, \mathbf{x}_k$ equal to the columns of M numbered $i_1, \dots, i_k$, respectively. The identity (18) in this case is immediate (the left-hand side is a diagonal matrix). To derive our main result, we only need the $(n, k)=(3, 2)$ case of (18), which can also be verified directly.

If M is an $n\times n$ matrix, $\mathbf{x}\in\mathbb{R}^n$, and $\mathbf{z}\in\mathbb{R}^{n+1}$, let $\langle\mathbf{z};M,\mathbf{x}\rangle$ denote the determinant of the $(n+1)\times(n+1)$-matrix with the first row $\mathbf{z}$ and the remaining rows consisting of the matrix M with $\mathbf{x}$ upended as the last column. We note that

$$\det\begin{pmatrix}\langle\mathbf{z};M,\mathbf{x}\rangle & \langle\mathbf{z};M,\mathbf{x}'\rangle\\ \langle\mathbf{z}';M,\mathbf{x}\rangle & \langle\mathbf{z}';M,\mathbf{x}'\rangle\end{pmatrix}=\langle M\rangle\Bigg(\sum_{1\le i<j\le n}\langle M;\mathbf{x},\mathbf{x}'\rangle_{i,j}\big(z_iz_j'-z_jz_i'\big)\\ -\sum_{1\le i\le n}\langle M;\mathbf{x}-\mathbf{x}'\rangle_i\big(z_iz_{n+1}'-z_{n+1}z_i'\big)\Bigg) \tag{19}$$

for all $\mathbf{x},\mathbf{x}'\in\mathbb{R}^n$ and $\mathbf{z},\mathbf{z}'\in\mathbb{R}^{n+1}$, for the following reason. Since the subspace of nondegenerate (invertible) matrices is dense, it is sufficient to verify (19) under the assumption that the rows of M span $\mathbb{R}^n$. Since both sides are linear and anti-symmetric in the inputs $\mathbf{z}$ and $\mathbf{z}'$, it is sufficient to verify (19) only with $\mathbf{z}$ and $\mathbf{z}'$ equal to different coordinate vectors $\mathbf{e}_i$ in $\mathbb{R}^{n+1}$. Since

$$\langle\mathbf{e}_i;M,\mathbf{x}\rangle=\begin{cases}(-1)^{n-1}\langle M;\mathbf{x}\rangle_i, & \text{if } i\le n;\\ (-1)^n\langle M\rangle, & \text{if } i=n+1;\end{cases}$$

the identity (19) with $\mathbf{z}$ or $\mathbf{z}'$ equal to $\mathbf{e}_{n+1}$ is immediate, and that with $(\mathbf{z},\mathbf{z}')=(\mathbf{e}_i,\mathbf{e}_j)$, where $1\le i<j\le n$, follows from the $k=2$ case of (18). To derive our main result, we only need the $n=3$ case of (19), which can also be verified directly from the $(n,k)=(3,2)$ case of (18).

If $\mathbf{z},\widetilde{\mathbf{z}}\in\mathbb{R}^{n+1}$ differ only by the last entry and the last entry of $\mathbf{z}'$ is 0, (19) gives

$$\begin{aligned}\det\begin{pmatrix}\langle\mathbf{z};M,\mathbf{x}\rangle & \langle\widetilde{\mathbf{z}};M,\mathbf{x}'\rangle\\ \langle\mathbf{z}';M,\mathbf{x}\rangle & \langle\mathbf{z}';M,\mathbf{x}'\rangle\end{pmatrix}&=\frac{\tilde z_{n+1}}{z_{n+1}}\det\begin{pmatrix}\langle\mathbf{z};M,\mathbf{x}\rangle & \langle\mathbf{z};M,\frac{z_{n+1}}{\tilde z_{n+1}}\mathbf{x}'\rangle\\ \langle\mathbf{z}';M,\mathbf{x}\rangle & \langle\mathbf{z}';M,\frac{z_{n+1}}{\tilde z_{n+1}}\mathbf{x}'\rangle\end{pmatrix}\\ &=\frac{\tilde z_{n+1}}{z_{n+1}}\langle M\rangle\Bigg(\sum_{1\le i<j\le n}\langle M;\mathbf{x},\frac{z_{n+1}}{\tilde z_{n+1}}\mathbf{x}'\rangle_{i,j}\big(z_iz_j'-z_jz_i'\big)+\sum_{1\le i\le n}\langle M;\mathbf{x}-\frac{z_{n+1}}{\tilde z_{n+1}}\mathbf{x}'\rangle_i\,z_{n+1}z_i'\Bigg)\\ &=\langle M\rangle\Bigg(\sum_{1\le i<j\le n}\langle M;\mathbf{x},\mathbf{x}'\rangle_{i,j}\big(z_iz_j'-z_jz_i'\big)+\sum_{1\le i\le n}\langle M;\tilde z_{n+1}\mathbf{x}-z_{n+1}\mathbf{x}'\rangle_i\,z_i'\Bigg).\end{aligned} \tag{20}$$

The three equalities above are valid if $z_{n+1},\tilde z_{n+1}\neq 0$, but the first and the last expressions in (20) are equal for all $z_{n+1},\tilde z_{n+1}$ (again by continuity).

Going back to our setting, for each $\ell=1,\ldots,4$, denote

- by M_ℓ the matrix in (4) with the ℓ-th row and the fourth column removed,
- by $\mathbf{z}_\ell$ and $\mathbf{z}_\ell'$ the ℓ-th row of the matrix in $D(\mathbf{p},\mathbf{q},\mathbf{1}_4)$ with $q_\ell=0$ and the q_ℓ-partial of the ℓ-th row of the same matrix, respectively (the latter is a constant function in $\mathbf{q}$),
- by $\widetilde{\mathbf{z}}_\ell$ and $\widetilde{\mathbf{z}}_\ell'$ the analogues for the matrix in $D(\mathbf{p},\mathbf{q},\mathbf{S}_X)$, and
- by $\mathbf{x}_\ell'\in\mathbb{R}^3$ the column vector $\mathbf{S}_X$ with the ℓ-th entry removed.

By equation (13),

$$D(\mathbf{p},\mathbf{q},\mathbf{1}_4)^2\cdot\frac{\partial s_X}{\partial q_\ell}=\det\begin{pmatrix}\langle\mathbf{z}_\ell;M_\ell,\mathbf{1}_3\rangle & \langle\widetilde{\mathbf{z}}_\ell;M_\ell,\mathbf{x}_\ell'\rangle\\ \langle\mathbf{z}_\ell';M_\ell,\mathbf{1}_3\rangle & \langle\widetilde{\mathbf{z}}_\ell';M_\ell,\mathbf{x}_\ell'\rangle\end{pmatrix}, \tag{21}$$

where $\mathbf{1}_3 = (1, 1, 1)$. The row vectors $\mathbf{z}_\ell$ and $\tilde{\mathbf{z}}_\ell$ differ only by the last entry (except for $\ell = 1$ when they are equal), while $\mathbf{z}'_\ell = \tilde{\mathbf{z}}'_\ell$ with the last entry equal to 0. Thus, by (21) and (20),

$$D(\mathbf{p}, \mathbf{q}, \mathbf{1}_4)^2 \cdot \frac{\partial s_X}{\partial q_\ell} = \langle M_\ell \rangle \Bigg(\sum_{1 \le i < j \le 3} \langle M_\ell; \mathbf{1}_3, \mathbf{x}'_\ell \rangle_{i,j} \left(z_{\ell;i} z'_{\ell;j} - z_{\ell;j} z'_{\ell;i} \right) + \sum_{1 \le i \le 3} \langle M_\ell; \tilde{z}_{\ell;4} \mathbf{1}_3 - \mathbf{x}'_\ell \rangle_i \, z'_{\ell;i} \Bigg), \tag{22}$$

since $z_{\ell;4} = 1$. Since each variable q_i (and p_i as well) appears in a single row of M, each of the two factors on the right-hand side above is at most linear in each of these variables.

References

[1] Adami, Christoph and Arend Hintze (2013), Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature Communications* 4, doi:10.1038/ncomms3193.

[2] Akin, Ethan (2013), Good Strategies for the Iterated Prisoner's Dilemma. *arXiv:1211.0969v2.*

[3] Axelrod, Robert (1984), *The Evolution of Cooperation.* Basic Books, New York, NY, USA.

[4] Axelrod, Robert and Douglas Dion (1988), The further evolution of cooperation. *Science*, 242, 1385–1390.

[5] Axelrod, Robert and William D. Hamilton (1981), The evolution of cooperation. *Science*, 211, 1390–1396.

[6] Bhaskar, V., George J. Mailath, and Stephen Morris (2008), Purification in the infinitely repeated prisoner's dilemma. *Review of Economic Dynamics*, 11, 515528.

[7] Bhaskar, V. and Ichiro Obara (2002), Belief-based equilibria in the repeated prisoner's dilemma with private monitoring. *Journal of Economic Theory*, 102, 40–69.

[8] Dawkins, Richard (1976), *The Selfish Gene.* Oxford University Press, Oxford, UK.

[9] Ely, Jeffrey C. and Juuso Välimäki (2002), A robust folk theorem for the prisoner's dilemma. *Journal of Economic Theory*, 102, 84–105.

[10] Hilbe, Christian, Martin A. Nowak, and Karl Sigmund (2013), Evolution of extortion in Iterated Prisoner's Dilemma games. *Proceedings of the National Academy of Sciences*, 110(17), 6913–6918.

[11] Hilbe, Christian, Martin A. Nowak, and Arne Traulsen (2013), Adaptive Dynamics of Extortion and Compliance. *PLoS ONE* 8(11): e77886.

[12] Kendall, Graham, Xin Yao, and Siang Yew Chong (2007), *The Iterated Prisoners' Dilemma: 20 Years On.* World Scientific, River Edge, NJ, USA.

[13] Nowak, Martin A. (2006), Five rules for the evolution of cooperation. *Science*, 314, 1560–1563.

[14] Nowak, Martin A. and Karl Sigmund (1993), A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364, 56–58.

[15] Papadimitriou, Christos H. and Mihalis Yannakakis (1994), On complexity as bounded rationality. *Proceedings of the twenty-sixth annual ACM Symposium on Theory of Computing*, 726–733.

[16] Poundstone, William (1992), *Prisoner's Dilemma.* Doubleday, New York, NY, USA.

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

[18] Roberts, Kevin (1985), Cartel behavior and adverse selection. *Journal of Industrial Economics*, 33, 401–413.

[19] Rubinstein, Ariel (1986), Finite automata play the repeated prisoner's dilemma. *Journal of Economic Theory*, 39, 83–96.

[20] Stewart, Alexander J. and Joshua B. Plotkin (2012), Extortion and cooperation in the prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 109, 10134–10135.

[21] Stewart, Alexander J. and Joshua B. Plotkin (2013), From extortion to generosity, evolution in the Iterated Prisoner's Dilemma. *Proceedings of the National Academy of Sciences*, doi: 10.1073/pnas.1306246110.

[22] Szolnoki, Attila and Matjaž Perc (2014), Evolution of extortion in structured populations. *Physical Review E* 89, 022804.