Reactive strategies with longer memory

Nikoleta E. Glynatsi, Ethan Akin, Martin Nowak, Christian Hilbe

1 Formal Model

We consider infinitely repeated games among two players, player 1 and player 2. Each round, they engage in the donation game with payoff matrix

$$\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}. \tag{1}$$

Here b and c denote the benefit and the cost of cooperation, respectively. We assume $b > c > 0$ throughout. Therefore, payoff matrix (1) is a special case of the prisoner's dilemma with payoff matrix,

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix}, \tag{2}$$

where $T > R > S > P$ and $2R > T + S$. Here, R is the reward payoff of mutual cooperation, T is the temptation to defect payoff, S is the sucker's payoff, and P is the punishment payoff for mutual defection.

We assume the repeated game is played infinitely many rounds, and that the players' decisions only depend on the outcome of the previous n rounds. To this end, an *n-history for player $i \in \{1, 2\}$* is a string $h^i = (a^i_{-n}, \ldots, a^i_{-1}) \in \{C, D\}^n$ where an entry a^i_{-k} corresponds to player i's action k rounds ago. Let H^i denote the space of all n-histories for player i. Set H^i contains $|H^i| = 2^n$ elements. A pair $h = (h^1, h^2)$ is called an *n-history of the game*. We use $H = H^1 \times H^2$ to denote the space of all such histories which contains $|H| = 2^{2n}$ elements.

A *memory-n* strategy is a vector $\mathbf{m} = (m_h)_{h \in H} \in [0, 1]^{2n}$. Each entry m_h corresponds to the player's cooperation probability in the next round, depending on the outcome of the previous n rounds. One special case of memory$-n$ strategies are the round$-k-$repeat strategies for some $1 \le k \le n$. Player 1 uses a *round-k-repeat strategy* $\mathbf{m}^{k-\text{Rep}}$ if in any given round, the player chooses the same action as k rounds ago. That is, if the game's n-history is such that,

$$\begin{cases} m_h^{k-\text{Rep}} = 1, & \text{if } a^1_{-k} = C \\ m_h^{k-\text{Rep}} = 0, & \text{if } a^1_{-k} = D. \end{cases}$$

Two additional special cases of memory$-n$ strategies that we will be discussing in this work are, reactive$-n$ and self-reactive-n strategies. A *reactive$-n$ strategy* for player 1 is a vector $\mathbf{p} = (p_h)_{h \in H^2} \in [0, 1]^n$. Each entry p_h corresponds to the player's cooperation probability in the next round, based on the co-player's actions in the previous n rounds. Therefore, reactive-n strategies exclusively rely on the co-player's n-history, independent of the focal player's own actions. On the other hand, *self-reactive-n* strategies only consider the focal player's own n-history, and ignore the co-player's. Formally, a self-reactive-n strategy for player 1 is a vector $\tilde{\mathbf{p}} = (\tilde{p}_h)_{h \in H^1} \in [0,1]^n$. Each entry $\tilde{p}_h$ corresponds to the player's cooperation probability in the next, depending on the player's own actions in the previous n rounds. From hereon, we will use the notations $\mathbf{m}, \mathbf{p},$ and $\tilde{\mathbf{p}}$ to denote memory-n, reactive-n, and self-reactive-n strategies.

Let players 1 and 2 use memory$-n$ strategies $\mathbf{m^1}$ and $\mathbf{m^2}$. Then one can represent the interaction as a Markov chain with possible states $h \in H$. Assume that the present round is given by $h = (h^1, h^2)$, the probability that one round later $\tilde{h}$ is observed is given by the product,

$$M_{h,\tilde{h}} = \prod_{i=1}^{2} x^i$$

where,

$$x^i = \begin{cases} m_h^i & \text{if } \tilde{\alpha}_{-1}^i = C \text{ and } \tilde{\alpha}_{-t}^i = \alpha_{-t+1}^i \text{ for all other } \tilde{\alpha}_{-t}^i \\ 1 - m_h^i & \text{if } \tilde{\alpha}_{-1}^i = D \text{ and } \tilde{\alpha}_{-t}^i = \alpha_{-t+1}^i \text{ for all other } \tilde{\alpha}_{-t}^i \\ 0 & \text{if } \tilde{\alpha}_{-t}^i \neq \alpha_{-t+1}^i \text{ for some } 2 \leq t \leq n. \end{cases}$$

Let $\mathbf{v} = (v_h)_{h \in H}$ be an invariant distribution of this Markov chain for which $\mathbf{v} = \mathbf{v} \cdot M$.

1.1 An Extension of Akin's Lemma

The work of [Akin, 2016] focuses on memory-one strategies. In the case of $n = 1$, a memory-one strategy is represented by the vector $\mathbf{m} = (m_{CC}, m_{CD}, m_{DC}, m_{DD})$. Let player 1 use the memory-one strategy $\mathbf{m}$ when interacting with player 2. This leads to a sequence of distributions $\{\mathbf{v}^t, t = 1, 2, ...\}$ with $\mathbf{v}^t$ representing the distribution over the states in the t^{th} round of the game. Let $\mathbf{v} = (v_{CC}, v_{CD}, v_{DC}, v_{DD})$ be an associated stationary distribution of the interaction. In his work Akin shows that,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \mathbf{v}^t \cdot (\mathbf{m} - \mathbf{m}^{1-\text{Rep}}) = 0, \text{ and therefore } \mathbf{v} \cdot (\mathbf{m} - \mathbf{m}^{1-\text{Rep}}) = 0. \tag{3}$$

With the same method as in [Akin, 2016], one can derive a generalized version of his result. Namely the generalized version is given by Lemma 1.1.

Lemma 1.1 (Generalized Akin Lemma)**.** Let player 1 use a memory-n strategy, and let player 2 use any arbitrary strategy. The interaction between the two players leads to a sequence of distributions $\{\mathbf{v}^t, k = 1, 2, ...\}$ with $\mathbf{v}^t$ representing the distribution over the states in the t^{th} round of the game. Let $\mathbf{v}$ be an associated stationary distribution of the interaction. Then for each k with $1 \leq k \leq n$, the invariant distribution $\mathbf{v}$ satisfies the following relationship,

$$\mathbf{v} \cdot (\mathbf{m} - \mathbf{m}^{k-\text{Rep}}) = \sum_{h \in H} v_h (m_h - m_h^{k-\text{Rep}}) = 0. \tag{4}$$

Proof. Let the probability that player 1 cooperated in the n^{th} round be denoted as v_{C}^{n}. Let v_{C}^{n} be defined as the probability that player 1 played C, $k\,(1 \leq k \leq n)$ rounds ago. Then,

$$v_{\mathrm{C}}^{n} = \sum_{h \in H} y_h, \qquad \text{where} \qquad y_h = \begin{cases} u_h & \text{if } \alpha_{-k}^{1} = C \\ 0 & \text{if } \alpha_{-k}^{1} = D. \end{cases}$$

Equivalently,

$$v_{\mathrm{C}}^{n} = \mathbf{v}^{n} \cdot \mathbf{m}^{k-\mathrm{Rep}}.$$

The probability that player 1 cooperates in the $(n+1)^{th}$ round, denoted by $v_{\mathrm{C}}^{n+1} = \mathbf{v}^{n} \cdot \mathbf{m}$. Hence,

$$v_{\mathrm{C}}^{n+1} - v_{\mathrm{C}}^{n} = \mathbf{v}^{n} \cdot \mathbf{m} - \mathbf{v}^{n} \cdot \mathbf{m}^{k-\mathrm{Rep}} = \mathbf{v}^{n} \cdot (\mathbf{m} - \mathbf{m}^{k-\mathrm{Rep}}).$$

This implies,

$$\sum_{t=1}^{n} \mathbf{v}^{t} \cdot (\mathbf{m} - \mathbf{m}^{k-\mathrm{Rep}}) = \sum_{t=1}^{n} v_{\mathrm{C}}^{t+1} - v_{\mathrm{C}}^{t} \quad \Rightarrow \quad \sum_{t=1}^{n} \mathbf{v}^{t} \cdot (\mathbf{m} - \mathbf{m}^{k-\mathrm{Rep}}) = v_{\mathrm{C}}^{n+1} - v_{\mathrm{C}}^{1}. \tag{5}$$

As the right side has absolute value at most 1,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \mathbf{v}^{t} \cdot (\mathbf{m} - \mathbf{m}^{k-\mathrm{Rep}}) = 0. \tag{6}$$

□

The intuition for this result is that $\mathbf{v} \cdot \mathbf{m}$ and all $\mathbf{v} \cdot \mathbf{m}^{k-\mathrm{Rep}}$ are just different, but equivalent, expressions for player 1's average cooperation rate. More specifically, $\mathbf{m}^{1-\mathrm{Rep}} = \mathbf{m}^{2-\mathrm{Rep}} = \cdots = \mathbf{m}^{n-\mathrm{Rep}}$ correspond to the intuition that it does not matter which of the past n rounds we use to define cooperation rate.

1.2 Payoffs and Further Definitions

The invariant distribution $\mathbf{v}$ can also be used to define the long term payoffs of the players'. Let $\mathbf{S}^{k} = (S_h^{k})_{h \in H}$ denote the vector that returns for each h the one-shot payoff that player 1 obtained k rounds ago,

$$S_h^{k} = \begin{cases} b - c & \text{if } a_{-k}^{1} = C \text{ and } a_{-k}^{2} = C \\ -c & \text{if } a_{-k}^{1} = C \text{ and } a_{-k}^{2} = D \\ b & \text{if } a_{-k}^{1} = D \text{ and } a_{-k}^{2} = C \\ 0 & \text{if } a_{-k}^{1} = D \text{ and } a_{-k}^{2} = D \end{cases} \tag{7}$$

Then we can define player 1's repeated-game payoff $s_{\mathbf{m^1},\mathbf{m^2}}$ as

$$s_{\mathbf{m^1},\mathbf{m^2}} = \mathbf{v} \cdot \mathbf{S}^{1} = \mathbf{v} \cdot \mathbf{S}^{2} = \ldots = \mathbf{v} \cdot \mathbf{S}^{n}. \tag{8}$$

The equalities $\mathbf{v} \cdot \mathbf{S}^1 = \ldots = \mathbf{v} \cdot \mathbf{S}^n$ correspond to the intuition that it does not matter which of the past n rounds we use to define average payoffs. This is an immediate result of Lemma 1.1. The payoffs of the players depend on both players' cooperations, and since their cooperation can be defined as having occurred in any of the last n turns, the payoffs can also be expressed analogously. The payoff $s_{\mathbf{m^2},\mathbf{m^1}}$ of player 2 can be defined in a similar manner.

Let's provide definitions for some additional terms that will be used in this manuscript.

Definition 1.1 (Nash Strategies)**.** A strategy $\mathbf{m}$ for player 1, is a *Nash strategy* if,

$$s_{\mathbf{m}',\mathbf{m}} \leq s_{\mathbf{m},\mathbf{m}} \; \forall \; \mathbf{m}'. \tag{9}$$

Definition 1.2 (Nice Strategies)**.** A player's strategy is *nice*, if the player is never the first to defect. A nice strategy against itself receives the mutual cooperation payoff, $(b - c)$.

Definition 1.3 (Partner Strategies)**.** A *partner strategy* is a strategy which is both nice and Nash.

Partners strategies are of interest because they are strategies that strive to achieve the mutual cooperation payoff of $(b - c)$ with their co-player. However, if the co-player doesn't cooperate, they are prepared to penalize them with lower payoffs. Partner strategies, by definition, are best responses to themselves [Hilbe et al., 2015]. All partner strategies are Nash strategies, but not all Nash strategies are partner strategies.

To check whether a strategy $\mathbf{m}$ is Nash, one has to check Eq. (9) against all possible memory-n strategies. However the works of [McAvoy and Nowak, 2019, McAvoy et al., 2022] have shown that for $n = 1$ it is sufficient to check against all pure memory-one strategies. Here we generalize this result to any n.

Lemma 1.2 (Monotinic Payoffs)**.** ...

Lemma 1.3 (Sufficiency of Pure Strategies)**.** ...

2 Tit For Tat and Generous Tit For Tat

Based on Lemma 1.1, we can derive a theory of zero-determinant strategies analogous to the case of memory-one strategies. In the following, we say a memory-n strategy $\mathbf{m^1}$ is a zero-determinant strategy if there are k_1, k_2, k_3 and α, β, γ such that $\mathbf{m^1}$ can be written as

$$\mathbf{m^1} = \alpha \mathbf{S}^{k_1} + \beta \tilde{\mathbf{S}}^{k_2} + \gamma \mathbf{1} + \mathbf{m}^{k_3-\text{Rep}}, \tag{10}$$

where $\mathbf{1}$ is the vector for which every entry is 1. By Akin's Lemma and the definition of payoffs,

$$0 = \mathbf{v} \cdot (\mathbf{m^1} - \mathbf{m}^{k_3-\text{Rep}}) = \mathbf{v} \cdot (\alpha \mathbf{S}^{k_1} + \beta \tilde{\mathbf{S}}^{k_2} + \gamma \mathbf{1}) = \alpha s_{\mathbf{m^1},\mathbf{m^2}} + \beta s_{\mathbf{m^2},\mathbf{m^1}} + \gamma. \tag{11}$$

That is, payoffs satisfy a linear relationship.

One interesting special case arises if $k_1 = k_2 = k_3 =: k$ and $\alpha = -\beta = 1/(b+c)$ and $\gamma = 0$. In that case, the formula (10) yields the strategy

$$m_h = \begin{cases} 1 & \text{if } \; a^2_{-k} = C \\ 0 & \text{if } \; a^2_{-k} = D \end{cases} \tag{12}$$

That is, this strategy implements Tit-for-Tat (for $k = 1$) or delayed versions thereof (for $k > 1$). These strategies are partners strategies that also satisfy a stronger relationship. By Eq. (11), the enforced payoff relationship is $s_{\mathbf{m^1},\mathbf{m^2}} = s_{\mathbf{m^2},\mathbf{m^1}}$.

Another interesting special case arises if $k_1 = k_2 = k_3 =: k$ and $\alpha = 0$, $\beta = -1/b$, $\gamma = 1 - c/b$. In that case Eq. (10) yields the strategy

$$m_h = \begin{cases} 1 & \text{if } a^2_{-k} = C \\ 1 - c/b & \text{if } a^2_{-k} = D \end{cases} \tag{13}$$

That is, the generated strategy is GTFT (if $k = 1$), or delayed versions thereof (for $k > 1$). By Eq. (11), the enforced payoff relationship is $s_{\mathbf{m^2},\mathbf{m^1}} = b - c$. In particular, these strategies are partner strategies.

3 Reactive Partner Strategies

We assume in the following, that 1 adopts reactive$-n$ strategy $\mathbf{p}$, and player 2 adopts an arbitrary memory-n strategy $\mathbf{m}$. We define the following marginal distributions with respect to the possible n-histories of player 2,

$$v^2_h = \sum_{h^1 \in H^1} v_{(h^1,h)}. \tag{14}$$

These entries describe how often we observe player 2 to choose actions h^2, in n consecutive rounds (irrespective of the actions of player 1). Note that,

$$\sum_{h \in H^2} v^2_h = 1. \tag{15}$$

Let $\mathbf{p^{k-\text{Rep}}}$ be a reactive round-k-repeat strategy. Then the cooperation rate of player 2, denoted as $\rho_{\mathbf{m}}$, is given by,

$$\rho_{\mathbf{m}} = \sum_{h \in H^2} v^q_h \cdot p^{1-\text{Rep}}_h \;=\; \sum_{h \in H^2} v^q_h \cdot p^{2-\text{Rep}}_h \;=\; \ldots \;=\; \sum_{h \in H^2} v^q_h \cdot p^{n-\text{Rep}}_h. \tag{16}$$

The cooperating rate of player 1 can also be expressed in terms of v^2_h as,

$$\rho_{\mathbf{p}} \;=\; \sum_{h \in H^2} v^2_h \cdot p_h. \tag{17}$$

Because we consider simple donation games, we note that these two quantities, $\rho_{\mathbf{m}}$ *and* $\rho_{\mathbf{p}}$, are sufficient to define the payoffs of the two players,

$$\begin{aligned} s_{\mathbf{p},\mathbf{m}} &= b\,\rho_{\mathbf{m}} - c\,\rho_{\mathbf{p}} \\ s_{\mathbf{m},\mathbf{q}} &= b\,\rho_{\mathbf{p}} - c\,\rho_{\mathbf{m}}. \end{aligned} \tag{18}$$

3.1 Sufficiency of Self reactive strategies

Press and Dyson [2012] discussed the case where one player uses a memory-one strategy and the other player employs a longer memory strategy. They demonstrated that the payoff of the player with the longer memory is exactly the same as if the player had employed a specific shorter-memory strategy, disregarding any history beyond what is shared with the short-memory player. Here we show a result that follows a similar intuition: if there is a part of history that one player does not observe, then the co-player gains nothing by considering the history not shared with the reactive player.

Lemma 3.1. Let $\mathbf{p}$ be a reactive$-n$ strategy for player 1. Then, for any memory$-n$ strategy $\mathbf{m}$ used by player 2, player 1's score is exactly the same as if 2 had played a specific self-reactive memory-n strategy $\tilde{\mathbf{p}}$.

Proof. ... □

Lemma 1.3 states that for a memory-n strategy to be Nash, then condition (9) has to hold against all pure memory-n strategies. This also applies to the case where the player adopts a reactive-n strategy. However, for the case of a reactive player here we show a stronger result. More specifically, combining Lemma 1.3 and Lemma 3.1 leads to the focal result,

Lemma 3.2. A reactive-n strategy $\mathbf{p}$ for player 1, is a *Nash strategy* if,

$$s_{\tilde{\mathbf{p}},\mathbf{p}} \leq s_{\mathbf{p},\mathbf{p}} \; \forall \; \tilde{p} \in \tilde{P}. \tag{19}$$

where,

$$\tilde{P} = \{(\tilde{p}_h) : \tilde{p}_h \in \{0,1\} \; \forall \; h \in H^2\}.$$

Thus, $\tilde{P}$ is the set of all pure self-reactive-n strategies.

Lemma 3.2 implies that the space of strategies we need to check against is even more constrained in the case of reactive strategies. This has a huge implication on the computational complexity of finding Nash strategies. In particular, the number of strategies one has to check against is reduced from $2^{2^{2n}}$ to 2^{2n}.

Another result of Lemma 3.1 is that we can retrieve the marginal distributions of the co-player's actions (Eq (14)) without having to consider the transition matrix M. For now, one has to calculate the transition matrix M for two given players and calculate the stationary distribution of this $2^{2n} \times 2^{2n}$ matrix. However, since the co-player is playing a self-reactive strategy then the co-player's action only rely on his/her actions, and thus one can model this as a Markov process with states $h^2 \in H^2$ and a transition matrix $\tilde{M}$. Let $h^2 = ((a^2_{-n}, \dots, a^2_{-1}))$ be the state in the current round. The probability that in the next turn $\tilde{h}^2 = ((\tilde{a}^2_{-n}, \dots, \tilde{a}^2_{-1}))$ is observed is given by,

$$\tilde{M}_{h^2,\tilde{h}^2} = \begin{cases} \tilde{p}_{h^2} & \text{if } \tilde{\alpha}^2_{-1} = C \text{ and } \tilde{\alpha}^2_{-t} = \alpha^2_{-t+1} \text{ for all other } \tilde{\alpha}^2_{-t} \\ 1 - \tilde{p}_{h^2} & \text{if } \tilde{\alpha}^2_{-1} = D \text{ and } \tilde{\alpha}^2_{-t} = \alpha^2_{-t+1} \text{ for all other } \tilde{\alpha}^2_{-t} \\ 0 & \text{if } \tilde{\alpha}^2_{-t} \neq \alpha^2_{-t+1} \text{ for some } 2 \leq t \leq n. \end{cases}$$

The stationary distribution of this Markov chain denoted as $\tilde{\mathbf{v}}$ has the property that,

$$\tilde{u}_h = u_h^2.$$

This results that the payoffs of the players can now be calculated more efficiently.

3.2 Reactive-Two Partner Strategies

In this section, we focus on the case of $n = 2$. Reactive-two strategies are denoted as a vector $\mathbf{p} = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$ where p_{CC} is the probability of cooperating in this turn when the co-player cooperated in the last 2 turns, p_{CD} is the probability of cooperating given that the co-player cooperated in the second to last turn and defected in the last, and so forth. A nice reactive-two strategy is represented by the vector $\mathbf{p} = (1, p_{CD}, p_{DC}, p_{DD})$.

Theorem 3.3 ("Reactive-Two Partner Strategies")**.** A nice reactive-two strategy $\mathbf{p}$, is a partner strategy if and only if, the strategy entries satisfy the conditions:

$$p_{DD} < 1 - \frac{c}{b} \quad and \quad \frac{p_{CD} + p_{DC}}{2} < 1 - \frac{1}{2} \cdot \frac{c}{b}. \tag{20}$$

There are two independent proffs of Theorem 3.3. The first prove is in line with the work of [Akin, 2016], and the second one relies on the sufficiency of self-reactive strategies. We discuss both proofs in the Appendix A.

3.3 Reactive-Three Partner Strategies

In this section, we focus on the case of $n = 3$. Reactive-three strategies are denoted as a vector

$$\mathbf{p} = (p_{CCC}, p_{CCD}, p_{CDC}, p_{CDD}, p_{DCC}, p_{DCD}, p_{DDC}, p_{DDD})$$

where p_{CCC} is the probability of cooperating in round t when the co-player cooperates in the last 3 rounds, p_{CCD} is the probability of cooperating given that the co-player cooperated in the third and second to last rounds and defected in the last, and so forth. A nice reactive-three strategy has $p_{CCC} = 1$.

Theorem 3.4 ("Reactive-Three Partner Strategies")**.** A nice reactive-three strategy $\mathbf{p}$, is a partner strategy if and only if, the strategy entries satisfy the conditions:

$$\begin{aligned}
\frac{p_{CCD} + p_{CDC} + p_{DCC}}{3} &< 1 - \frac{1}{3} \cdot \frac{c}{b} \\
\frac{p_{CDD} + p_{DCD} + p_{DDC}}{3} &< 1 - \frac{2}{3} \cdot \frac{c}{b} \\
p_{DDD} &< 1 - \frac{c}{b} \\
\frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{4} &< 1 - \frac{1}{2} \cdot \frac{c}{b} \\
\frac{p_{CDC} + p_{DCD}}{2} &< 1 - \frac{1}{2} \cdot \frac{c}{b}
\end{aligned} \tag{21}$$

Once again, there are two independent proves of Theorem 3.4, and we discuss both proofs in the Appendix B.

3.4 Reactive Counting Strategies

A special case of reactive strategies is reactive counting strategies. These are strategies that respond to the co-player's actions, but they do not distinguish between when cooperations/defections occurred; they solely consider the count of cooperations in the last n turns. A reactive-n counting strategy is represented by a vector $\mathbf{r} = (r_i)_{i \in \{n, n-1, \ldots, 0\}}$, where the entry r_i indicates the probability of cooperating given that the co-player cooperated i times in the last n turns.

Reactive-two counting strategies are denoted by the vector $\mathbf{r} = (r_2, r_1, r_0)$. We can characterise partner strategies among the reactive-two counting strategies by setting $r_2 = 1$, and $p_{CD} = p_{DC} = r_1$ and $p_{DD} = r_0$ in conditions (20). This gives us the following result.

Corollary 3.4.1. A nice reactive-two counting strategy $\mathbf{r} = (1, r_1, r_0)$ is a partner strategy if and only if,

$$r_1 < 1 - \frac{1}{2} \cdot \frac{c}{b} \quad \textit{and} \quad r_0 < 1 - \frac{c}{b}. \tag{22}$$

Reactive-three counting strategies are denoted by the vector $\mathbf{r} = (r_3, r_2, r_1, r_0)$. We can characterise partner strategies among reactive-three counting strategies by setting $r_3 = 1$, and $p_{CCD} = p_{CDC} = p_{DCC} = r_2, p_{DCD} = p_{DDC} = p_{CDD} = r_1$ and $p_{DDD} = r_0$ in conditions (21). This gives us the following result.

Corollary 3.4.2. A nice reactive-three counting strategy $\mathbf{r} = (1, r_2, r_1, r_0)$ is a partner strategy if and only if,

$$r_2 < 1 - \frac{1}{3} \cdot \frac{c}{b}, \quad r_1 < 1 - \frac{2}{3} \cdot \frac{c}{b} \quad \textit{and} \quad r_0 < 1 - \frac{c}{b}. \tag{23}$$

In the case of counting reactive strategies, we generalize to the case of n.

Corollary 3.4.3 ("Reactive-Counting Partner Strategies")**.** A nice reactive-n counting strategy $\mathbf{r} = (r_i)_{i \in \{n, n-1, \ldots, 0\}}$, is a partner strategy if and only if:

$$r_{n-k} < 1 - \frac{k}{n} \cdot \frac{c}{b}, \text{ for } k \in \{1, 2, \ldots, n\}. \tag{24}$$

4 Prisoner's Dilemma

So far we have focused on a special case of the Prisoner's Dilemma, the donation game. In this section we show that the results of Sections 3.2 and 3.3 can be generalized for the iterated Prisoner's Dilemma.

We assume that player 1 uses a reactive-n strategy $\mathbf{p}$, and that the co-player uses a self-reactive strategy $\tilde{\mathbf{p}}$. The interaction can be model as a Markov proceed with a stationary distribution $\mathbf{v}$ as we have already discussed. The long-term payoff of player 2 is now given by,

$$\mathbf{s}_{\tilde{\mathbf{p}}, \mathbf{p}} = a_R \cdot R + a_S \cdot S + a_T \cdot T + a_P \cdot P, \quad \textit{where}$$

$$
\begin{aligned}
a_R &= \sum_{h \in H^2} u_h^2 \cdot p_h \cdot \tilde{p}_h, \\
a_S &= \sum_{h \in H^2} u_h^2 \cdot (1 - p_h) \cdot \tilde{p}_h, \\
a_T &= \sum_{h \in H^2} u_h^2 \cdot p_h \cdot (1 - \tilde{p}_h), \\
a_P &= \sum_{h \in H^2} u_h^2 \cdot (1 - p_h) \cdot (1 - \tilde{p}_h).
\end{aligned}
$$

The payoff for player 2 can be expressed analogously.

For the case of reactive-two strategies.

Corollary 4.0.1. A nice reactive-two strategy $\mathbf{p}$, is a partner strategy if and only if, the strategy entries satisfy the conditions:

$$
\begin{aligned}
(P - T)\, p_{DD} &< P - R, \\
(T - P)\,(p_{CD} + p_{DC}) + (R - S)\, p_{DD} &< 3R - S - 2\,P, \\
(T - P)\, p_{CD} + (R - S)\,(p_{CD} + p_{DD}) &< 4\,R - 2\,S - P - T, \\
(R - S)\,(p_{CD} + p_{DC}) &< 3R - 2S - T, \\
(R - S)\ p_{CD} + (T - P)\ p_{DC} &< 2R - S - P
\end{aligned}
$$

For the case of reactive-three strategies.

Corollary 4.0.2. A nice reactive-three strategy $\mathbf{p}$, is a partner strategy if and only if, the strategy entries satisfy the conditions:

$$
\begin{aligned}
p_{DDD}(R-S)+(T-P)(p_{CDD}+p_{DCD}+p_{DDC}) &< 4R-3P-S\\
p_{CDC}(T-P)+p_{DCD}(R-S) &< 2R-P-S\\
-P(p_{DDD}-1)+Tp_{DDD} &< R\\
(T-P)(p_{CCD}+p_{CDD}+p_{DDC})+(R-S)(p_{CDC}+p_{DCC}+p_{DCD}+p_{DDD}) &< 8R-3P-4S-T\\
p_{DCC}(T-P)+(R-S)(p_{CCD}+p_{CDC}) &< 3R-P-2S\\
(T-P)(p_{CCD}+p_{DCC}+p_{DDC})+(R-S)(p_{CDC}+p_{CDD}+p_{DCD}) &< 6R-3P-3S\\
(T-P)(p_{CCD}+p_{DDC})+(R-S)(p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD}) &< 7R-2P-4S-T\\
(T-P)(p_{CCD}+p_{CDD}+p_{DCC})+(R-S)(p_{DDC}+p_{DDD}) &< 5R-3P-2S\\
p_{CDD}(R-S)+(T-P)(p_{DCD}+p_{DDC}) &< 3R-2P-S\\
p_{CCD}(T-P)+(R-S)(p_{CDD}+p_{DCC}+p_{DDC}) &< 5R-P-3S-T\\
(T-P)(p_{CCD}+p_{DCC})+(R-S)(p_{CDD}+p_{DDC}) &< 4R-2P-2S\\
(T-P)(p_{CDC}+p_{DCD})+(R-S)(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC}) &< 7R-2P-4S-T\\
(T-P)(p_{CDC}+p_{CDD}+p_{DCD})+(R-S)(p_{CCD}+p_{DCC}+p_{DDC}+p_{DDD}) &< 8R-3P-4S-T\\
(T-P)(p_{CDC}+p_{DCC}+p_{DCD})+(R-S)(p_{CCD}+p_{CDD}+p_{DDC}) &< 6R-3P-3S\\
(T-P)(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC})+(R-S)(p_{CDC}+p_{DCD}+p_{DDD}) &< 7R-4P-3S\\
(R-S)(p_{CCD}+p_{CDC}+p_{DCC}) &< 4R-3S-T\\
(T-P)(p_{CCD}+p_{CDD})+(R-S)(p_{DCC}+p_{DDC}+p_{DDD}) &< 6R-2P-3S-T\\
(T-P)(p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD})+(R-S)(p_{CCD}+p_{DDC}+p_{DDD}) &< 7R-4P-3S
\end{aligned}
$$

A Proofs for Theorem 3.3

A.1 Approach based on Akin's Generalised Lemma

Suppose player 1 adopts a reactive-two strategy $\mathbf{p}=(p_{CC},p_{CD},p_{DC},p_{DD})$. Moreover, suppose player 2 adopts an arbitrary memory-2 strategy $\mathbf{m}$. Let $\mathbf{v}=(v_h)_{h\in H}$ be an invariant distribution of the game between the two players.

The cooperation rate of player 2 given by 16 in the case of $n=2$ is given by,

$$\rho_{\mathbf{m}} := v^2_{CC}+v^2_{CD} = v^2_{CC}+v^2_{DC}. \tag{25}$$

We can use this equality to conclude that

$$v^2_{CD} = v^2_{DC}. \tag{26}$$

Moreover the cooperation rate of player 1 based on Eq. 17 is given by,

$$\begin{aligned}\rho_{\mathbf{p}} &= v_{CC}^2\,p_{CC} + v_{CD}^2\,p_{CD} + v_{DC}^2\,p_{DC} + v_{DD}^2\,p_{DD} \\ &= v_{CC}^2\,p_{CC} + v_{CD}^2\,(p_{CD}{+}p_{DC}) + v_{DD}^2\,p_{DD}.\end{aligned} \tag{27}$$

Here, the second equality is due to Eq. (26).

Proof. ($\Rightarrow$)A reactive-two strategy $\mathbf{p} = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$ can only be a Nash equilibrium if *no* other strategy yields a larger payoff, in particular neither AllD nor the Alternator strategy must yield a larger payoff, where

$$\text{AllD} = (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0) \text{ and Alternator} = (0,0,1,1,0,0,1,1,0,0,1,1,0,0,1,1).$$

Thus, $\mathbf{p}$ can only form a Nash equilibrium if

$$\pi(\text{AllD}, \mathbf{p}) \leq b{-}c \quad \text{ and } \quad \pi(\text{Alternator}, \mathbf{p}) \leq b{-}c,$$

or equivalently, if

$$p_{DD} \leq 1 - \frac{c}{b} \quad \text{ and } \quad p_{CD} + p_{DC} \leq 1 + \frac{b{-}c}{c}. \tag{28}$$

($\Leftarrow$) Now, suppose player 2 has some strategy $\mathbf{m}$ such that $s_{\mathbf{m},\mathbf{p}} > b{-}c$. It follows that

$$\begin{aligned} 0 &< s_{\mathbf{m},\mathbf{p}} - (b{-}c) \\ &\overset{Eq.\ (18)}{=} b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b{-}c) \\ &\overset{Eqs.\ (25),(27),(15)}{=} b\left(v_{CC}^2 p_{CC} + v_{CD}^2 (p_{CD}{+}p_{DC}) + v_{DD}^2 p_{DD}\right) - c\left(v_{CC}^2 + v_{CD}^2\right) - (b{-}c)\left(v_{CC}^2 + 2v_{CD}^2 + v_{DD}^2\right) \\ &= v_{CC}^2\, b\,(p_{CC} - 1) + v_{CD}^2\Big(b(p_{CD}{+}p_{DC}) + c - 2b\Big) + v_{DD}^2\Big(bp_{DD} - (b{-}c)\Big). \end{aligned} \tag{29}$$

Condition (29) can hold only if,

$$b\,(p_{CD}{+}p_{DC}) + c - 2b > 0, \qquad b\,p_{DD} - (b{-}c) > 0. \tag{30}$$

Thus, Eq. (28) reassures that $\mathbf{p}$ is Nash strategy, and given that $p_{CC} = 1$, it is a partner strategy. □

A.2 Approach based on Self-Reactive Sufficiency Lemma

Suppose player 1 adopts a nice reactive-two strategy $\mathbf{p}=(1, p_{CD}, p_{DC}, p_{DD})$. For $\mathbf{p}$ to be a Nash strategy,

$$s_{\tilde{\mathbf{p}},\mathbf{p}} \leq (b-c), \tag{31}$$

must hold against all $\tilde{\mathbf{p}} \in \tilde{P}$, where $\tilde{P}$ is the set of all pure self-reactive-two strategies. In the case of $n=2$, the set contains 16 strategies.

Proof. Suppose player 1 plays a nice reactive-two strategy $\mathbf{p} = (1, p_{CD}, p_{DC}, p_{DD})$, and suppose the co-player 2 plays a pure self-reactive-two strategy $\tilde{\mathbf{p}}$. The possible payoffs for $\tilde{\mathbf{p}} \in \{\tilde{\mathbf{p}}^0, \ldots, \tilde{\mathbf{p}}^{16}\}$ are:

$$\begin{array}{rcll}
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = & b \cdot p_{DD} & for \;\; i \in & \{0,2,4,6,8,10,12,14\} \\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = & \frac{b\cdot(p_{CD}+p_{DC}+p_{DD})}{3} - \frac{1}{3}\cdot c & for \;\; i \in & \{1,9\} \\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = & \frac{b\cdot(p_{CD}+p_{DC}+p_{DD}+1)}{4} - \frac{1}{2}\cdot c & for \;\; i \in & \{3\} \\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = & \frac{b\cdot(p_{CD}+p_{DC})}{2} - \frac{1}{2}\cdot c & for \;\; i \in & \{4,5,12,13\} \\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = & \frac{b\cdot(p_{CD}+p_{DC}+1)}{3} - \frac{2}{3}\cdot c & for \;\; i \in & \{6,7\} \\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = & b-c & for \;\; i \in & \{8,9,10,11,12,13,14,15\}
\end{array}$$

Setting the payoff expressions of $s_{\tilde{\mathbf{p}}^i,\mathbf{p}}$ to smaller or equal to $(b-c)$ we get the following unique conditions,

$$p_{DD} \leq 1 - \frac{c}{b} \tag{32}$$

$$\frac{p_{CD}+p_{DC}}{2} \leq 1 - \frac{1}{2}\cdot\frac{c}{b} \tag{33}$$

$$\frac{p_{CD}+p_{DC}+p_{DD}}{3} \leq 1 - \frac{2}{3}\cdot\frac{c}{b} \tag{34}$$

Notice that only conditions (32) and (33) are necessary.

□

B Proofs for Theorem 3.4

B.1 Approach based on Akin's Generalised Lemma

Suppose player 1 adopts a reactive-three strategy $\mathbf{p}$, and suppose player 2 adopts an arbitrary memory-three strategy $\mathbf{m}$. Let $\mathbf{v} = (v_h)_{h\in H}$ be an invariant distribution of the game between the two players.

The average cooperation rate $\rho_{\mathbf{m}}$ of player 2 (Eq. 16) for $n = 3$ is given by,

$$\rho_{\mathbf{m}} := v^2_{CCC}+v^2_{CCD}+v^2_{DCC}+v^2_{DCD} = v^2_{CCC}+v^2_{DCC}+v^2_{CDC}+v^2_{DDC} = v^2_{CCC}+v^2_{CCD}+v^2_{CDC}+v^2_{CDD}. \quad (35)$$

We can use this equality to conclude that

$$\begin{aligned} v^2_{CCD} &= v^2_{DCC} && (36)\\ v^2_{DDC} &= v^2_{CDD} && (37)\\ v^2_{CCD}+v^2_{DCD} = v^2_{CDC}+v^2_{DDC} \Rightarrow v^2_{CCD} &= v^2_{CDC}+v^2_{CDD}-v^2_{DCD} && (38) \end{aligned}$$

The average cooperation rate of 1's (Eq. (17)) for $n = 3$ is given by,

$$\begin{aligned} \rho_{\mathbf{p}} &= v^2_{CCC}\,p_{CCC} + v^2_{CCD}\,p_{CCD} + v^2_{CDC}\,p_{CDC} + v^2_{CDD}\,p_{CDD} + v^2_{DCD}\,p_{DCD} + \\ &\quad + v^2_{DDC}\,p_{DDC} + v^2_{DDD}\,p_{DDD} \\ &\overset{Eqs.\ (36),(37)}{=} v^2_{CCC}\,p_{CCC} + v^2_{CCD}\,(p_{CCD}+p_{DCC}) + v^2_{CDC}\,p_{CDC} + v^2_{CDD}\,(p_{CDD}+p_{DDC}) + \\ &\quad v^2_{DCD}\,p_{DCD} + v^2_{DDD}\,p_{DDD} \end{aligned} \quad (39)$$

Proof. ($\Rightarrow$) A reactive-three strategy $\mathbf{p}$ can only be a Nash equilibrium if *no* other strategy yields a larger payoff, in particular neither AllD nor the following self-reactive-three strategies,

$$\begin{aligned} \tilde{\mathbf{p}}^{\mathbf{15}} &= (0,0,0,0,1,1,1,1)\\ \tilde{\mathbf{p}}^{\mathbf{17}} &= (0,0,0,1,0,0,0,1)\\ \tilde{\mathbf{p}}^{\mathbf{51}} &= (0,0,1,1,0,0,1,1)\\ \tilde{\mathbf{p}}^{\mathbf{119}} &= (0,1,1,1,0,1,1,1). \end{aligned}$$

The above strategies are alternating strategies.For instance, $\tilde{\mathbf{p}}^{\mathbf{15}}$ and $\tilde{\mathbf{p}}^{\mathbf{51}}$ are delayed alternating strategies. $\tilde{\mathbf{p}}^{\mathbf{15}}$ cooperates if and only if defected three rounds ago, and $\tilde{\mathbf{p}}^{\mathbf{15}}$ cooperates after defecting 2 rounds ago. $\tilde{\mathbf{p}}^{\mathbf{17}}$ and $\tilde{\mathbf{p}}^{\mathbf{119}}$ alternate between cooperating and defecting after given sequences occur. Namely, $\tilde{\mathbf{p}}^{\mathbf{17}}$ cooperates after DD sequence has occurred, and $\tilde{\mathbf{p}}^{\mathbf{119}}$ defects after CCC sequence has occurred.

$\mathbf{p}$ can only form a Nash equilibrium if

$$\pi(\text{AllD}, \mathbf{p}) \leq b - c \quad and \quad \pi(\tilde{\mathbf{p}}^i, \mathbf{p}) \leq b - c \text{ for } i \in \{15, 17, 51, 102\}.$$

or equivalently, if

$$
\begin{aligned}
\frac{p_{CCD}+p_{CDC}+p_{DCC}}{3} &< 1-\frac{1}{3}\cdot\frac{c}{b}\\
\frac{p_{CDD}+p_{DCD}+p_{DDC}}{3} &< 1-\frac{2}{3}\cdot\frac{c}{b}\\
p_{DDD} &< 1-\frac{c}{b}\\
\frac{p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC}}{4} &< 1-\frac{1}{2}\cdot\frac{c}{b}\\
\frac{p_{CDC}+p_{DCD}}{2} &< 1-\frac{1}{2}\cdot\frac{c}{b}
\end{aligned}
\tag{40}
$$

($\Leftarrow$) Now, suppose player 2 has some strategy $\mathbf{m}$ such that $s_{\mathbf{m},\mathbf{p}} > b-c$. It follows that

$$
\begin{aligned}
0 \quad &\leq \quad s_{\mathbf{m},\mathbf{p}} - (b-c)\\
&\overset{Eq.\ (18)}{=} \quad b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b-c)\\
&\overset{Eqs.\ (39),(15)}{=} \quad b\left(v^2_{CCC}p_{CCC}+v^2_{CCD}(p_{CCD}+p_{DCC})+v^2_{CDC}p_{CDC}+v^2_{DDC}(p_{CDD}+p_{DDC})+v^2_{DCD}p_{DCD}+v^2_{DDD}p_{DDD}\right)\\
&\qquad -c\left(v^2_{CCC}+2v^2_{CCD}+v^2_{DCD}\right) - (b-c)\left(v^2_{CCC}+2v^2_{CCD}+v^2_{CDC}+2v^2_{DDC}+v^2_{DCD}+v^2_{DDD}\right)\\
&= \quad b\,v^2_{CCC}(p_{CCC}-1) + v^2_{CCD}(b\,(p_{CCD}+p_{DCC}-2)) + v^2_{CDC}(b\,(p_{CDC}-1)+c)+\\
&\qquad v^2_{CDD}(b\,(p_{CDD}+p_{DDC}-2)+2\,c) + v^2_{DCD}(b\,(p_{DCD}-1)) + v^2_{DDD}(b\,(p_{DDD}-1)+c)\\
&\overset{Eq.\ (38)}{=} \quad b\,v^2_{CCC}(p_{CCC}-1) + v^2_{DDD}(b\,(p_{DDD}-1)+c) + v^2_{CDC}(b\,(p_{CCD}+p_{DCC}+p_{CDC}-3)+c)+\\
&\qquad v^2_{CDD}(b\,(p_{CDD}+p_{DDC}+p_{CCD}+p_{DCC}-4)+2\,c) + v^2_{DCD}(b\,(p_{DCD}-1) - b\,(p_{CCD}+p_{DCC})-2)
\end{aligned}
\tag{41}
$$

Condition (41) holds only for,

$$
b\,(p_{DDD}-1)+c<0, \quad b\,(p_{CCD}+p_{DCC}+p_{CDC}-3)+c
$$

$$
b\,(p_{CDD}+p_{DDC}+p_{CCD}+p_{DCC}-4)+2\,c<0 \Rightarrow -b\,(p_{CCD}+p_{DCC}-2) > b\,(p_{CDD}+p_{DDC}-2)+2\,c
$$

$$
b\,(p_{DCD}-1) - b\,(p_{CCD}+p_{DCC})-2<0 \Rightarrow b\,(p_{DCD}+p_{CDD}+p_{DDC}-3)+2\,c<0.
$$

Thus, conditions Eq. (40) reassure that $\mathbf{p}$ is Nash strategy, and given that $p_{CC}=1$, it is a partner strategy. □

B.2 Approach based on Self-Reactive Sufficiency Lemma

Consider all the pure self-reactive-three strategies. There is a total of 256 such strategies. The payoff expression for each of these strategies against a nice reactive-three strategies can be calculated explicitly. We use these expressions to obtain the conditions for partner strategies similar to the previous section.

Proof. The payoff expressions for a nice reactive-three strategy $\mathbf{p}$ against all pure self-reactive-three strategies are as follows,

$$\begin{aligned}
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= b\, p_{DDD} && for \;\; i \in \{0,2,4,6,\ldots,250,252,254\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CDD}+p_{DCD}+p_{DDC}+p_{DDD})}{4} - \tfrac{1}{4}\cdot c && for \;\; i \in \{1,9,33,41,65,73,97,105,129,137,161,\\
&&& \quad 169,193,201,225,233\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC}+p_{DDD})}{5} - \tfrac{2}{5}\cdot c && for \;\; i \in \{3,7,35,39,131,135,163,167\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CDC}+p_{DCD})}{2} - \tfrac{1}{2}\cdot c && for \;\; i \in \{4-7,12-15,20-23,28-31,68-71,\\
&&& \quad 76-79,84-87,92-95,132-135,\\
&&& \quad 140-143,148-151,156-159,\\
&&& \quad 196-199,204-207,212-215,220-223\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC}+p_{DDD}+1)}{6} - \tfrac{1}{2}\cdot c && for \;\; i \in \{11,15,43,47\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CDD}+p_{DCD}+p_{DDC})}{3} - \tfrac{1}{3}\cdot c && for \;\; i \in \{16,17,24,25,48,49,56,57,80,81,88,\\
&&& \quad 89,112,113,120,121,144,145,152,153,\\
&&& \quad 176,177,184,185,208,209,216,217,\\
&&& \quad 240,241,248,249\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC})}{4} - \tfrac{1}{2}\cdot c && for \;\; i \in \{18,19,22,23,50,51,54,55,146,147,\\
&&& \quad 150,151,178,179,182,183\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC}+1)}{5} - \tfrac{3}{5}\cdot c && for \;\; i \in \{26,27,30,31,58,59,62,63\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CCD}+p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD}+p_{DDC}+p_{DDD})}{7} - \tfrac{3}{7}\cdot c && for \;\; i \in \{37,67,165,195\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CCD}+p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD}+p_{DDC}+p_{DDD}+1)}{8} - \tfrac{1}{2}\cdot c && for \;\; i \in \{45,75\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CCD}+p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD}+p_{DDC})}{6} - \tfrac{1}{2}\cdot c && for \;\; i \in \{52,53,82,83,180,181,210,211\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CCD}+p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD}+p_{DDC}+1)}{7} - \tfrac{4}{7}\cdot c && for \;\; i \in \{60,61,90,91\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CCD}+p_{CDC}+p_{DCC})}{3} - \tfrac{2}{3}\cdot c && for \;\; i \in \{96-103,112-119,224-231,240-247\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{b\cdot(p_{CCD}+p_{CDC}+p_{DCC}+1)}{4} - \tfrac{3}{4}\cdot c && for \;\; i \in \{104-111,120-127\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= (b-c) && for \;\; i \in \{128,129,130,\ldots,255\}
\end{aligned} \tag{42}$$

Setting these to smaller or equal than the mutual cooperation payoff $(b-c)$ give the following ten conditions,

$$p_{DDD} \le 1-\frac{c}{b}, \quad \frac{p_{CDC}+p_{DCD}}{2} \le 1-\frac{1}{2}\cdot\frac{c}{b}, \quad \frac{p_{CDD}+p_{DCD}+p_{DDC}}{3} \le 1-\frac{2}{3}\cdot\frac{c}{b}, \tag{43}$$

$$\frac{p_{CCD}+p_{CDC}+p_{DCC}}{3} \le 1-\frac{1}{3}\cdot\frac{c}{b}, \quad \frac{p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC}}{4} \le 1-\frac{1}{2}\cdot\frac{c}{b} \tag{44}$$

$$\frac{p_{CDD}+p_{DCD}+p_{DDC}+p_{DDD}}{4} \le 1-\frac{3}{4}\cdot\frac{c}{b}, \tag{45}$$

$$\frac{p_{CCD}+p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD}+p_{DDC}+p_{DDD}}{7} \le 1-\frac{4}{7}\cdot\frac{c}{b}, \tag{46}$$

$$\frac{p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC}+p_{DDD}}{5} \le 1-\frac{3}{5}\cdot\frac{c}{b}, \tag{47}$$

$$\frac{p_{CCD}+p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD}+p_{DDC}}{6} \le 1-\frac{1}{2}\cdot\frac{c}{b} \tag{48}$$

Notice that only the conditions of Eq. (43) and (44) are necessary. The remaining conditions can be derived from the sums of conditions in Eq. (43) and (44). $\square$

C Proof of Corollary 3.4.3

To prove corollary 3.4.3 we need to introduce some additional notation. We introduce the vector $\mathbf{w} = (w_i)_{i \in \{0,1,\ldots,n\}}$, where the entry w_i is the probability that in the long term outcome the co-player cooperates i times.

An element of $\mathbf{w}$ is the sum of one or more of the marginal distribution $u^2_{h^2}$ for $h^2 \in H^2$. Namely let,

$$H_i^2 = \{h^2 : |a_C(h^2)| = i \quad \forall \quad h^2 \in H^2\}, \text{ where}$$

$$a_C(h^2) = \{a^2_{-t} : a^2_{-t} = C \quad \forall \quad a^2_{-t} \in h^2\}.$$

Then we define w_i as,

$$w_i = \sum_{h \in H_i^2} v_h.$$

The cooperation rate of the reactive player is given by,

$$\rho_{\mathbf{p}} = \sum_{i=0}^{n} r_i \cdot w_i. \tag{49}$$

The co-player can use any self-reactive-n strategy, and thus the co-player differentiates between when the last cooperation/defection occurred. However, we can still express the co-player's cooperation rate as a function of w_i. More specifically, the co-player's cooperation rate is,

$$\rho_{\tilde{\mathbf{p}}} = \sum_{i=0}^{n} \frac{i}{n} \cdot w_i. \tag{50}$$

We will also define a set of $i-$ repeat strategies. Assume the set of self reactive strategies $A = \{\mathbf{A^0}, \mathbf{A^1}, \ldots, \mathbf{A^n}\}$, where.

The payoff of an alternating self-reactive-n against a counting-reactive$-n$ $\mathbf{r}$ is given by,

$$s_{\mathbf{A^i},\mathbf{r}} = b \cdot r_i - \frac{i}{n} \cdot c \quad for \quad i \in [0, n]. \tag{51}$$

The intuition behind Eq. (51) is that in the long term, the strategies end up in a state where $\mathbf{A^i}$ has cooperated i times in the last n turns. Thus, the co-player will cooperate and provide the benefit b with a probability r_i, while in return, the alternating strategy has cooperated $\frac{i}{n}$ times and pays the cost.

With this we have all the required tools to prove the following theorem.

Proof. ($\Rightarrow$) As we have already discussed previously, a strategy can only be a Nash equilibrium if the payoff of the co-player does not exceed $(b-c)$. Therefore, for p to be a Nash equilibrium against each strategy in set A (for $i \in [0,n]$),

$$s_{\mathbf{A^i},\mathbf{r}} \leq b - c \tag{52}$$
$$b \cdot r_i - \frac{i}{n} \cdot c \leq b - c \tag{53}$$
$$r_i \leq 1 - \frac{i}{n} \cdot \frac{c}{b} \tag{54}$$

Now, suppose player q has some strategy $\mathbf{m}$ and player p has a reactive-counting strategy such that $s_{\mathbf{m},\mathbf{p}} > b$–$c$. It follows that

$$\begin{aligned} 0 &\leq s_{\mathbf{m},\mathbf{p}} - (b{-}c) \\ &\overset{Eq.\ (18)}{=} b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b{-}c) \\ &\overset{Eqs.\ (\mathbf{??}),(49),(50)}{=} b\sum_{k=0}^{n} r_{n-k} \cdot u_{n-k} - c\sum_{k=0}^{n} \frac{n-k}{n} \cdot u_{n-k} - (b-c)\sum_{k=0}^{n} u_{n-k} \\ &\quad u_n\Big(b\,(r_n - 1)\Big) + \sum_{k=1}^{n} u_{n-k}\Big(b\sum_{k=1}^{n} r_{n-k} - c\sum_{k=0}^{n-1} \frac{n-k}{n} - (b-c)\Big) \end{aligned} \tag{55}$$

This condition holds only if,

$$\Big(b\, r_{n-k} - c\,\frac{n-k}{n} - (b-c)\Big) < 0 \Rightarrow \tag{56}$$
$$b\,(r_{n-k} - 1) + (1 - \frac{n-k}{n})\, c < 0 \Rightarrow \tag{57}$$
$$r_{n-k} < 1 - \frac{n}{k} \cdot \frac{c}{b}. \tag{58}$$

for $k \in [0,n]$. Thus, any counting strategy that satisfies conditions (52) is Nash, and if it is nice, it's also a partner strategy. $\square$

D Proofs for Corollaries 4.0.1 and 4.0.2

D.1 Reactive-Two Partner Strategies

There are 16 pure-self reactive strategies in $n = 2$. We use calculate the explicit payoff expressions for each pure self-reactive strategy against a nice reactive-two strategy as given by Eq. (4). This gives the following payoff expressions:

$$
\begin{aligned}
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= P(1-p_{DD})+Tp_{DD} && for \quad i\in\{0,2,4,6,8,10,12,14\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{-P(p_{CD}+p_{DC}-2)+Rp_{DD}-S(p_{DD}-1)+T(p_{CD}+p_{DC})}{3} && for \quad i\in\{1,9\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{P(1-p_{CD})+R(p_{DC}+p_{DD})-S(p_{DC}+p_{DD}-2)+T(p_{CD}+1)}{4} && for \quad i\in\{3\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{P(1-p_{DC})+Rp_{CD}-S(p_{CD}-1)+Tp_{DC}}{2} && for \quad i\in\{4,5,12,13\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= \tfrac{R(p_{CD}+p_{DC})-S(p_{CD}+p_{DC}-2)+T}{3} && for \quad i\in\{6,7\}\\
s_{\tilde{\mathbf{p}}^i,\mathbf{p}} &= R && for \quad i\in\{8,9,10,11,12,13,14,15\}
\end{aligned}
$$

Setting the above expressions to $\leq R$ gives the following conditions,

$$
\begin{aligned}
(P-T)\,p_{DD} &< P-R,\\
(T-P)\,(p_{CD}+p_{DC})+(R-S)\,p_{DD} &< 3R+S-2\,P,\\
(T-P)\,p_{CD}+(R-S)\,(p_{CD}+p_{DC}) &< 4\,R-2\,S-P-T,\\
(R-S)\;p_{CD}+(T-P)\;p_{DC} &< 2R-S-P,\\
(R-S)\,(p_{CD}+p_{DC}) &< 3R-2S-T.
\end{aligned}
$$

D.2 Reactive-Three Partner Strategies

Previously as in the previous subsection we calculate the explicit payoff expressions for each $\tilde{\mathbf{p}} \in \tilde{P}$ against a nice reactive-three. The set of pure self-reactive strategies $\tilde{P}$ in $n = 3$ contains 256 elements. The expressions for each strategy are given below,

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CDD}+p_{DCD}+p_{DDC})+3\,P+(R-S)\,p_{DDD}+S}{4} \qquad for \quad i \in \{1,9,33,41,65,73,97,105,129,137,161,169,193,201,225,233\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,p_{CDC}+P+(R-S)\,p_{DCD}+S}{2} \qquad for \quad i \in \{4-7,12-15,20-23,28-31,68-71,76-79,84-87,92-95,132-135,140-143,148-151,156-159,196-199,204-207,212-215,220-223\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = -P\,(p_{DDD}-1)+Tp_{DDD} \qquad for \quad i \in \{0,2,4,\ldots,252,254\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CCD}+p_{CDD}+p_{DDC})+3\,P+(R-S)\,(p_{CDC}+p_{DCC}+p_{DCD}+p_{DDD})+4\,S+T}{8} \qquad for \quad i \in \{45\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,p_{DCC}+P+(R-S)\,(p_{CDC}+p_{CCD})+2\,S}{3} \qquad for \quad i \in \{96-103,112-119,224-231,240-247\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CCD}+p_{DCC}+p_{DDC})+3\,P+(R-S)\,(p_{CDC}+p_{CDD}+p_{DCD})+3\,S}{6} \qquad for \quad i \in \{52,53,180,181\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CCD}+p_{DDC})+2\,P+T+(R-S)\,(p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD})+4\,S}{7} \qquad for \quad i \in \{60,61\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CCD}+p_{CDD}+p_{DCC})+3\,P+(R-S)\,(p_{DDC}+p_{DDD})+2\,S}{5} \qquad for \quad i \in \{3,7,35,39,131,135,163,167\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{DCD}+p_{DDC})+2\,P+(R-S)\,p_{CDD}+S}{3} \qquad for \quad i \in \{16,17,24,25,48,49,56,57,80,81,88,89,112,113,120,121,144,145,152,153,176,177,184,185,208,209,216,217,240,241,248,249\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = R \qquad for \quad i \in \{128,129,\ldots,255\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,p_{CCD}+P+T+(R-S)\,(p_{CDD}+p_{DCC}+p_{DDC})+3S}{5} \qquad for \quad i \in \{26,27,30,31,58,59,62,63\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CCD}+p_{DCC})+2\,P+(R-S)\,(p_{CDD}+p_{DDC})+2\,S}{4} \qquad for \quad i \in \{18,19,22,23,50,51,54,55,146,147,150,151,178,179,182,183\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CDC}+p_{DCD})+2\,P+T+(R-S)\,(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC})+4\,S}{7} \qquad for \quad i \in \{90,91\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CDC}+p_{CDD}+p_{DCD})+3\,P+T+(R-S)\,(p_{CCD}+p_{DCC}+p_{DDC}+p_{DDD})+4\,S}{8} \qquad for \quad i \in \{75\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CDC}+p_{DCC}+p_{DCD})+3\,P+(R-S)\,(p_{CCD}+p_{CDD}+p_{DDC})+3\,S}{6} \qquad for \quad i \in \{82,83,210,211\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC})+4\,P+(R-S)\,(p_{CDC}+p_{DCD}+p_{DDD})+3\,S}{7} \qquad for \quad i \in \{37,165\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{T+(R-S)\,(p_{CCD}+p_{CDC}+p_{DCC})+3\,S}{4} \qquad for \quad i \in \{104-111,120-127\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-P)\,(p_{CCD}+p_{CDD})+2\,P+T+(R-S)\,(p_{DCC}+p_{DDC}+p_{DDD})+3\,S}{6} \qquad for \quad i \in \{11,15,43,47\}$$

$$s_{\tilde{\mathbf{p}}^i,\mathbf{p}} = \frac{(T-p)\,(p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD})+4\,P+(R-S)\,(p_{CCD}+p_{DDC}+p_{DDD})+3\,S}{7} \qquad for \quad i \in \{67,195\}$$

Setting the above expressions to $\leq R$ gives the following conditions,

$$
\begin{array}{rcc}
p_{DDD}(R-S)+(T-P)(p_{CDD}+p_{DCD}+p_{DDC}) & < & 4R-3P-S \\
p_{CDC}(T-P)+p_{DCD}(R-S) & < & 2R-P-S \\
-P(p_{DDD}-1)+Tp_{DDD} & < & R \\
(T-P)(p_{CCD}+p_{CDD}+p_{DDC})+(R-S)(p_{CDC}+p_{DCC}+p_{DCD}+p_{DDD}) & < & 8R-3P-4S-T \\
p_{DCC}(T-P)+(R-S)(p_{CCD}+p_{CDC}) & < & 3R-P-2S \\
(T-P)(p_{CCD}+p_{DCC}+p_{DDC})+(R-S)(p_{CDC}+p_{CDD}+p_{DCD}) & < & 6R-3P-3S \\
(T-P)(p_{CCD}+p_{DDC})+(R-S)(p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD}) & < & 7R-2P-4S-T \\
(T-P)(p_{CCD}+p_{CDD}+p_{DCC})+(R-S)(p_{DDC}+p_{DDD}) & < & 5R-3P-2S \\
p_{CDD}(R-S)+(T-P)(p_{DCD}+p_{DDC}) & < & 3R-2P-S \\
p_{CCD}(T-P)+(R-S)(p_{CDD}+p_{DCC}+p_{DDC}) & < & 5R-P-3S-T \\
(T-P)(p_{CCD}+p_{DCC})+(R-S)(p_{CDD}+p_{DDC}) & < & 4R-2P-2S \\
(T-P)(p_{CDC}+p_{DCD})+(R-S)(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC}) & < & 7R-2P-4S-T \\
(T-P)(p_{CDC}+p_{CDD}+p_{DCD})+(R-S)(p_{CCD}+p_{DCC}+p_{DDC}+p_{DDD}) & < & 8R-3P-4S-T \\
(T-P)(p_{CDC}+p_{DCC}+p_{DCD})+(R-S)(p_{CCD}+p_{CDD}+p_{DDC}) & < & 6R-3P-3S \\
(T-P)(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC})+(R-S)(p_{CDC}+p_{DCD}+p_{DDD}) & < & 7R-4P-3S \\
(R-S)(p_{CCD}+p_{CDC}+p_{DCC}) & < & 4R-3S-T \\
(T-P)(p_{CCD}+p_{CDD})+(R-S)(p_{DCC}+p_{DDC}+p_{DDD}) & < & 6R-2P-3S-T \\
(T-P)(p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD})+(R-S)(p_{CCD}+p_{DDC}+p_{DDD}) & < & 7R-4P-3S
\end{array}
$$

References

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

C. Hilbe, A. Traulsen, and K. Sigmund. Partners or rivals? strategies for the iterated prisoner's dilemma. *Games and economic behavior*, 92:41–52, 2015.

A. McAvoy and M. A. Nowak. Reactive learning strategies for iterated games. *Proceedings of the Royal Society A*, 475(2223):20180819, 2019.

A. McAvoy, J. Kates-Harbeck, K. Chatterjee, and C. Hilbe. Evolutionary instability of selfish learning in repeated games. *PNAS nexus*, 1(4):pgac141, 2022.

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