Reactive strategies with longer memory

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

1 Model

We consider infinitely repeated games among two players, player 1 and player 2. Each round, they engage in the donation game (choosing to cooperate (C) or defect (D)) with a payoff matrix.

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \end{array} \tag{1}$$

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,

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} \tag{2}$$

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 that, in each round, players' decisions depend only on the outcome of the previous n rounds. An n-history for player $i \in \{1, 2\}$ is represented as $h^i = (a^i_{-n}, \ldots, a^i_{-1}) \in \{C, D\}^n$, where a^i_{-k} corresponds to player i's action k rounds ago. We denote the space of all n-histories for player i as H^i, which contains $|H^i| = 2^n$ elements. Note that we use the notation $-i$ to denote the co-player; for example, h^{-i} represents the n-history of the co-player.

A pair $h = (h^1, h^2)$ is referred to as 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.

Memory-n strategies. Memory-n strategies are represented by 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.

The interaction between two players using memory-n strategies, $\mathbf{m^1}$ and $\mathbf{m^2}$, can be modeled as a Markov chain with H possible states. Given that the current round is $h = (h^1, h^2)$, 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:

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

A transition matrix M of size $2^{2n} \times 2^{2n}$ is constructed to represent these probabilities.

Let $\mathbf{v} = (v_h)_{h \in H}$ denote the probability distribution over the states of the Markov chain. 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. A distribution $\mathbf{v}$ is considered a stationary distribution if $\mathbf{v} = \mathbf{v} \cdot M$.

Reactive-n strategies. In this work, we will discuss special cases of memory-n strategies, including reactive-n and self-reactive-n strategies. A *reactive-n strategy* is represented by a vector $\mathbf{p} = (p_{h^{-i}})_{h^{-i} \in H^{-i}} \in [0, 1]^n$. Each entry $p_{h^{-i}}$ corresponds to the player's cooperation probability in the next round, based on the co-player's actions in the previous n rounds. Reactive-n strategies rely exclusively on the co-player's n-history, independent of the focal player's own actions.

Self-reactive-n strategies. The final set of strategies we will discuss in this work are self-reactive-n strategies. A *self-reactive-n strategy* only considers the focal player's own n-history, ignoring the co-player's history. Formally, a self-reactive-n strategy is represented by a vector $\tilde{\mathbf{p}} = (\tilde{p}_{h^i})_{h^i \in H^i} \in [0, 1]^n$. Each entry $\tilde{p}_{h^i}$ corresponds to the player's cooperation probability in the next round, depending solely on the player's own actions in the previous n rounds.

Round$-k-$repeat strategies. One special case of strategies is the round-k-repeat strategies for some $1 \leq k \leq n$. A *round-k-repeat strategy* is a strategy where, in any given round, the player chooses the same action as they did in the round that occurred k rounds ago. For example, in the case of memory-n strategies, a round-k-repeat strategy is denoted as $\mathbf{m}^{k-\text{Rep}}$ if the game's n-history is such that,

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

From this point forward, we will use the notations $\mathbf{m}$, $\mathbf{p}$, and $\tilde{\mathbf{p}}$ to denote memory-n, reactive-n, and self-reactive-n strategies, respectively. Moreover, we will use the notations $\mathbf{m}^{k-\text{Rep}}$, $\mathbf{p}^{k-\text{Rep}}$, and $\tilde{\mathbf{p}}^{k-\text{Rep}}$ to denote memory-n, reactive-n, and self-reactive-n round-k-repeat strategy, respectively.

An Extension of Akin's Lemma In the case of $n = 1$, a memory-1 strategy is represented by the vector $\mathbf{m} = (m_{CC}, m_{CD}, m_{DC}, m_{DD})$. In his work Akin [2016] 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 player 1 use a memory-1 strategy $\mathbf{m}$ and player 2 an arbitrary memory-n strategy. 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,

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

Equivalently,

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

Let k be fixed to $k = 1$ then,

$$v_{\text{C}}^n = \mathbf{v}^n \cdot \mathbf{m}^{1-\text{Rep}}.$$

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

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

This implies,

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

As the right side has absolute value at most 1,

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

Repeat for $1 < k \leq n$.

□

The intuition behind this result is that $\mathbf{v}\cdot\mathbf{m}$ and all $\mathbf{v}\cdot\mathbf{m}^{k-\mathrm{Rep}}$ are simply different but equivalent expressions for player 1's average cooperation rate. To be more specific, $\mathbf{m}^{1-\mathrm{Rep}} = \mathbf{m}^{2-\mathrm{Rep}} = \cdots = \mathbf{m}^{n-\mathrm{Rep}}$ correspond to the idea that it doesn't matter which of the past n rounds we use to define the cooperation rate.

Payoffs. Here, we define the long-term payoffs of players. Initially, we establish the payoffs of the players in a single round. Let a^i_{-k} and a^{-i}_{-k} denote the actions of the player and the co-player in the k-th round, respectively. Then, $\mathbf{S}^i_k = (S^k_h)_{h\in H}$ represents the vector that returns, for each h, the one-shot payoff obtained k rounds ago.

$$S^k_h = \begin{cases} b-c & \text{if } a^i_{-k} = C \text{ and } a^{-i}_{-k} = C \\ -c & \text{if } a^i_{-k} = C \text{ and } a^{-i}_{-k} = D \\ b & \text{if } a^i_{-k} = D \text{ and } a^{-i}_{-k} = C \\ 0 & \text{if } a^i_{-k} = D \text{ and } a^{-i}_{-k} = D \end{cases} \tag{7}$$

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

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

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.

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

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

$$s_{\mathbf{m^{-i}},\mathbf{m}} \leq s_{\mathbf{m^i},\mathbf{m^i}} \ \forall\ \mathbf{m^{-i}}. \tag{9}$$

Note that for a memory-n strategy, we only need to check the above condition for all mutant strategies of memory n. This is a result from the work of Press and Dyson [2012].

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.

2 Tit For Tat and Generous Tit For Tat across All Memory Length

Building upon Lemma 1.1, we can develop a theory of zero-determinant strategies within the class of memory-n strategies. In the following, we say a memory-n strategy $\mathbf{m}$ is a zero-determinant strategy if there are k_1,

k_2, k_3 and α, β, γ such that $\mathbf{m^i}$ can be written as

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

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

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

That is, payoffs satisfy a linear relationship. Thus, $\mathbf{m^i}$ is a zero-determinant strategy.

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

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

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

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

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

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 $s_{\mathbf{m^{-i}},\mathbf{m^i}} = b - c$. In particular, these strategies are partner strategies.

3 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. . . . □

NG: pip install christian_hilbe

NG: Could you please try to come up with the proof?

Several results arise from Lemma 3.1. For example, assume that player 1 is using a reactive-n strategy, and player 2 is using a memory-n strategy. If we wanted to verify that player 1's strategy is Nash, we would have to show that condition (9) holds for all memory-n mutant strategies. However, now we can only consider self-reactive-n mutant strategies for player 2.

From here on, when we are considering the case where player 1 is using a reactive-n strategy, we will assume that player 2 is using a self-reactive-n strategy.

A more efficient way to calculate payoffs. Assume that player 1 is playing a reactive-n strategy $\mathbf{p}$ and the co-player is playing an arbitrary strategy $\tilde{\mathbf{p}}$. To calculate the long-term payoffs of the two players, we need to compute the stationary distribution of the Markov chain M and then calculate the payoffs using Eq. (8). However, since the co-player is playing a self-reactive strategy, their actions only rely on their own actions. This can be modeled as a Markov process with H^2 states and a transition matrix $\tilde{M}$. Let $h^2 = ((a^2_{-n}, \ldots, a^2_{-1}))$ be the state in the current round. The probability that in the next turn $\tilde{h}^2 = ((\tilde{a}^2_{-n}, \ldots, \tilde{a}^2_{-1}))$ is observed is given by,

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

and let $\tilde{\mathbf{v}}$ be a stationary distribution of the Markov chain $\tilde{M}$. Now player 1's payoff ($s_{\tilde{\mathbf{p}},\mathbf{p}}$) for the general Prisoner's Dilemma is given by,

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

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

Player 2's payoff is calculated in a similar way. In the case of the simple donation games, it is sufficient to define the payoffs of the two players based on their cooperation rates. More specifically, we can define the payoffs of the two players as,

$$\begin{aligned} s_{\mathbf{p},\tilde{\mathbf{p}}} &= b\,\rho_{\tilde{\mathbf{p}}} - c\,\rho_{\mathbf{p}} \\ s_{\tilde{\mathbf{p}},\mathbf{q}} &= b\,\rho_{\mathbf{p}} - c\,\rho_{\tilde{\mathbf{p}}}. \end{aligned} \tag{14}$$

where,

$$\rho_{\tilde{\mathbf{p}}} = \tilde{\mathbf{v}} \cdot \tilde{\mathbf{p}}^{1-\text{Rep}}, \quad \textit{and} \tag{15}$$

$$\rho_{\mathbf{p}} = \sum_{h^2 \in H^2} \tilde{v}^{h^2} \cdot p_{h^2}. \tag{16}$$

Let's consider an example in the case of $n = 1$ to demonstrate the result. Assume that player 1 is playing the reactive-1 strategy $\mathbf{p} = (\frac{1}{2}, \frac{1}{3}, \frac{1}{2}, \frac{1}{3})$, and player 2 is playing the self-reactive-1 strategy $\tilde{\mathbf{p}} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, \frac{1}{2})$. The transition matrix M is given by,

$$M = \begin{bmatrix} \frac{1}{8} & \frac{3}{8} & \frac{1}{8} & \frac{3}{8} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{8} & \frac{3}{8} & \frac{1}{8} & \frac{3}{8} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

The stationary distribution is given by $\mathbf{v} = (\frac{4}{25}, \frac{6}{25}, \frac{6}{25}, \frac{9}{25})$. The payoffs for players 1 and 2 are equal and are equal to $s_{\mathbf{p},\tilde{\mathbf{p}}} = s_{\tilde{\mathbf{p}},\mathbf{q}} = \frac{2}{5}(b - c)$.

Now we repeat the exercise using $\tilde{M}$.

$$\tilde{M} = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

The stationary distribution is given by $\tilde{\mathbf{v}} = (\frac{2}{5}, \frac{3}{5})$, and the cooperation rates $\rho_{\tilde{\mathbf{p}}} = \frac{2}{5}$ and $\rho_{\mathbf{p}} = \frac{2}{5} \times \frac{1}{2} + \frac{3}{5} \times \frac{1}{3} = \frac{2}{5}$. Thus, based on Eq. (14), the payoffs are $s_{\mathbf{p},\tilde{\mathbf{p}}} = s_{\tilde{\mathbf{p}},\mathbf{q}} = \frac{2}{5}(b - c)$.

4 An algorithm for Nash Equilibria amongst Reactive Strategies

To predict which reactive-n strategies are partner strategies, we must characterize which nice reactive-n strategies are Nash equilibria. Determining whether a given strategy, $\mathbf{p}$, is a Nash equilibrium is not straightforward. In principle, this would involve comparing the payoff of $\mathbf{p}$ to the payoff of all possible self-reactive strategies, and there are infinitely many self-reactive strategies. However, here we demonstrate that we can reduce the space of mutant strategies to a finite set.

Initially, we will demonstrate that the payoff of a self-reactive strategy against a reactive strategy is linear in the self-reactive strategy entries. Secondly, we will show that because the payoff is linear, the best response self-reactive strategy to the reactive strategy is a pure self-reactive strategy.

Lemma 4.1. A reactive-n strategy $\mathbf{p}$ for player 1, is a *Nash strategy* if, and only if, no pure self-reactive-n strategy can achieve a higher payoff against itself.

Lemma 4.1 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.

NG: Christian could you please take over this section? We have already discussed the proofs.

5 Reactive Partner Strategies

In this section, we will characterize the reactive-n partner strategies. We will present a series of results that are mainly proven based on Lemma 4.1. However, interestingly, a number of our results can also be proven by an independent proof based on the generalization of Akin (Lemma 1.1). We will present both proofs in the Appendices A and B.

5.1 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 5.1 ("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:

$$p_{DD} < 1 - \frac{c}{b} \quad \textit{and} \quad \frac{p_{CD} + p_{DC}}{2} < 1 - \frac{1}{2} \cdot \frac{c}{b}. \tag{17}$$

For proofs see Appendix sections A.1 and B.2.

5.2 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 5.2 ("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:

$$\begin{aligned}
\frac{p_{CCD} + p_{CDC} + p_{DCC}}{3} &< 1 - \frac{1}{3} \cdot \frac{c}{b} \\
\frac{p_{CDD} + p_{DCD} + p_{DDC}}{3} &< 1 - \frac{2}{3} \cdot \frac{c}{b} \\
p_{DDD} &< 1 - \frac{c}{b} \\
\frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{4} &< 1 - \frac{1}{2} \cdot \frac{c}{b} \\
\frac{p_{CDC} + p_{DCD}}{2} &< 1 - \frac{1}{2} \cdot \frac{c}{b}
\end{aligned} \tag{18}$$

For proofs see Appendix sections A.2 and B.3.

5.3 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 (17). This gives us the following result.

Corollary 5.2.1. A nice reactive-two counting strategy $\mathbf{r} = (1, r_1, r_0)$ is a partner strategy if and only if,

$$r_1 < 1 - \frac{1}{2} \cdot \frac{c}{b} \;\; \textit{and} \;\; r_0 < 1 - \frac{c}{b}. \tag{19}$$

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 (18). This gives us the following result.

Corollary 5.2.2. A nice reactive-three counting strategy $\mathbf{r} = (1, r_2, r_1, r_0)$ is a partner strategy if and only if,

$$r_2 < 1 - \frac{1}{3} \cdot \frac{c}{b}, \quad r_1 < 1 - \frac{2}{3} \cdot \frac{c}{b} \;\; \textit{and} \;\; r_0 < 1 - \frac{c}{b}. \tag{20}$$

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

Corollary 5.2.3 ("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:

$$r_{n-k} < 1 - \frac{k}{n} \cdot \frac{c}{b}, \text{ for } k \in \{1, 2, \ldots, n\}. \tag{21}$$

For proof see Appendix section C.

5.4 General 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 5.1 and 5.2 can be generalized for the iterated Prisoner's Dilemma. For the case of reactive-two strategies.

Corollary 5.2.4. A nice reactive-two strategy $\mathbf{p}$, is a partner strategy if and only if, the strategy entries satisfy the conditions:

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

For proof see Appendix section A.3.

For the case of reactive-three strategies.

Corollary 5.2.5. A nice reactive-three strategy $\mathbf{p}$, is a partner strategy if and only if, the strategy entries satisfy the conditions:

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

For proof see Appendix section A.4.

6 Evolutionary Simulations for Strong Selection

In the manuscript we present results on different values of the selection strength β. Here we present the results for the case of strong selection ($\beta = 100$). The results are qualitatively similar to the case of weak selection ($\beta = 1$). However, we note that.

A Proofs for Theorems Based on Pure Self-reactive Strategies Result

A.1 Proof of Theorem 5.1

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,

$$s_{\tilde{\mathbf{p}},\mathbf{p}} \leq (b-c), \tag{22}$$

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 $\tilde{\mathbf{p}} \in \{\tilde{\mathbf{p}}^0, \dots, \tilde{\mathbf{p}}^{16}\}$ are:

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

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,

$$p_{DD} \leq 1 - \frac{c}{b} \tag{23}$$
$$\frac{p_{CD}+p_{DC}}{2} \leq 1 - \frac{1}{2}\cdot\frac{c}{b} \tag{24}$$
$$\frac{p_{CD}+p_{DC}+p_{DD}}{3} \leq 1 - \frac{2}{3}\cdot\frac{c}{b} \tag{25}$$

Notice that only conditions (23) and (24) are necessary.

□

A.2 Proof of Theorem 5.2

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,\dots,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,\\
&&&&& 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,\\
&&&&& 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}} &= \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,\\
&&&&& 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}} &= \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,\\
&&&&& 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,\dots,255\}
\end{aligned}
\tag{26}
$$

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

$$p_{DDD} \leq 1 - \frac{c}{b}, \quad \frac{p_{CDC} + p_{DCD}}{2} \leq 1 - \frac{1}{2} \cdot \frac{c}{b}, \quad \frac{p_{CDD} + p_{DCD} + p_{DDC}}{3} \leq 1 - \frac{2}{3} \cdot \frac{c}{b}, \tag{27}$$

$$\frac{p_{CCD} + p_{CDC} + p_{DCC}}{3} \leq 1 - \frac{1}{3} \cdot \frac{c}{b}, \quad \frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{4} \leq 1 - \frac{1}{2} \cdot \frac{c}{b} \tag{28}$$

$$\frac{p_{CDD} + p_{DCD} + p_{DDC} + p_{DDD}}{4} \leq 1 - \frac{3}{4} \cdot \frac{c}{b}, \tag{29}$$

$$\frac{p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC} + p_{DDD}}{7} \leq 1 - \frac{4}{7} \cdot \frac{c}{b}, \tag{30}$$

$$\frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC} + p_{DDD}}{5} \leq 1 - \frac{3}{5} \cdot \frac{c}{b}, \tag{31}$$

$$\frac{p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC}}{6} \leq 1 - \frac{1}{2} \cdot \frac{c}{b} \tag{32}$$

Notice that only the conditions of Eq. (27) and (28) are necessary. The remaining conditions can be derived from the sums of conditions in Eq. (27) and (28). □

A.3 Proof of Corollary 5.2.4

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. (3). 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,

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

A.4 Proof of Corollary 5.2.5

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,

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

B Proofs for Theorems Based on Generalized Akin's Lemma

B.1 Further Notation

In this section, we introduce some additional notation that will be using in the following subsection to prove our theorems.

We once again assume the setup that player 1 adopts a reactive-n strategy $\mathbf{p}$, and player 2 adopts a self-reactive-n strategy $\tilde{\mathbf{p}}$. We define the following marginal distributions with respect to the possible n-histories of player 2:

$$v^2_{h^2} = \sum_{h^1 \in H^1} v_{(h^1,h^2)}. \tag{33}$$

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,

$$\sum_{h \in H^2} v_h^2 = 1. \tag{34}$$

Let $\mathbf{p}^{\mathbf{k}-\mathrm{Rep}}$ be a reactive round-k-repeat strategy. Then the cooperation rate of player 2, denoted as $\rho_{\tilde{\mathbf{p}}}$, and based on Lemma 1.1 is given by,

$$\rho_{\tilde{\mathbf{p}}} = \sum_{h^2 \in H^2} v_{h^2}^2 \cdot p_h^{1-\mathrm{Rep}} \; = \; \sum_{h^2 \in H^2} v_{h^2}^2 \cdot p_h^{2-\mathrm{Rep}} \; = \; \ldots \; = \sum_{h^2 \in H^2} v_{h^2}^2 \cdot p_h^{n-\mathrm{Rep}}. \tag{35}$$

Player's 1 cooperation rate can also be defined in a similar manner. However, here we define the cooperation rate of player 1 as,

$$\rho_{\mathbf{p}} \quad = \quad \sum_{h^2 \in H^2} v_{h^2}^2 \cdot p_{h^2}. \tag{36}$$

B.2 Proof of Theorem 5.1

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 35 in the case of $n = 2$ is given by,

$$\rho_{\mathbf{m}} := v_{CC}^2 + v_{CD}^2 = v_{CC}^2 + v_{DC}^2. \tag{37}$$

We can use this equality to conclude that

$$v_{CD}^2 = v_{DC}^2. \tag{38}$$

Moreover the cooperation rate of player 1 based on Eq. 36 is given by,

$$\begin{aligned} \rho_{\mathbf{p}} \quad &= \quad v_{CC}^2 \, p_{CC} + v_{CD}^2 \, p_{CD} + v_{DC}^2 \, p_{DC} + v_{DD}^2 \, p_{DD} \\ &= \quad v_{CC}^2 \, p_{CC} + v_{CD}^2 \, (p_{CD} + p_{DC}) + v_{DD}^2 \, p_{DD}. \end{aligned} \tag{39}$$

Here, the second equality is due to Eq. (38).

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

$$p_{DD} \leq 1 - \frac{c}{b} \quad \text{and} \quad p_{CD} + p_{DC} \leq 1 + \frac{b-c}{c}. \tag{40}$$

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

$$\begin{aligned}
0 &< s_{\mathbf{m},\mathbf{p}} - (b-c) \\
&\overset{Eq.\ (14)}{=} b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b-c) \\
&\overset{Eqs.\ (37),(39),(34)}{=} b\left(v^2_{CC}p_{CC} + v^2_{CD}(p_{CD}+p_{DC}) + v^2_{DD}p_{DD}\right) - c\left(v^2_{CC}+v^2_{CD}\right) - (b-c)\left(v^2_{CC}+2v^2_{CD}+v^2_{DD}\right) \\
&= v^2_{CC}\, b\,(p_{CC}-1) + v^2_{CD}\left(b(p_{CD}+p_{DC})+c-2b\right) + v^2_{DD}\left(bp_{DD} - (b-c)\right).
\end{aligned} \tag{41}$$

Condition (41) can hold only if,

$$b\,(p_{CD}+p_{DC})+c-2b > 0, \qquad b\,p_{DD} - (b-c) > 0. \tag{42}$$

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

B.3 Proof of Theorem 5.2

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. 35) for $n = 3$ is given by,

$$\rho_{\mathbf{m}} := v^2_{CCC} + v^2_{CCD} + v^2_{DCC} + v^2_{DCD} = v^2_{CCC} + v^2_{DCC} + v^2_{CDC} + v^2_{DDC} = v^2_{CCC} + v^2_{CCD} + v^2_{CDC} + v^2_{CDD}. \tag{43}$$

We can use this equality to conclude that

$$v^2_{CCD} = v^2_{DCC} \tag{44}$$

$$v^2_{DDC} = v^2_{CDD} \tag{45}$$

$$v^2_{CCD} + v^2_{DCD} = v^2_{CDC} + v^2_{DDC} \Rightarrow v^2_{CCD} = v^2_{CDC} + v^2_{CDD} - v^2_{DCD} \tag{46}$$

The average cooperation rate of 1's (Eq. (36)) for $n = 3$ is given by,

$$\begin{aligned} \rho_{\mathbf{p}} &= v^2_{CCC}\, p_{CCC} + v^2_{CCD}\, p_{CCD} + v^2_{CDC}\, p_{CDC} + v^2_{CDD}\, p_{CDD} + v^2_{DCD}\, p_{DCD} + \\ &\quad + v^2_{DDC}\, p_{DDC} + v^2_{DDD}\, p_{DDD} \\ &\overset{Eqs.\ (44),(45)}{=} v^2_{CCC}\, p_{CCC} + v^2_{CCD}\, (p_{CCD} + p_{DCC}) + v^2_{CDC}\, p_{CDC} + v^2_{CDD}\, (p_{CDD} + p_{DDC}) + \\ &\quad v^2_{DCD}\, p_{DCD} + v^2_{DDD}\, p_{DDD} \end{aligned} \tag{47}$$

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

$$\begin{aligned} \frac{p_{CCD} + p_{CDC} + p_{DCC}}{3} &< 1 - \frac{1}{3} \cdot \frac{c}{b} \\ \frac{p_{CDD} + p_{DCD} + p_{DDC}}{3} &< 1 - \frac{2}{3} \cdot \frac{c}{b} \\ p_{DDD} &< 1 - \frac{c}{b} \\ \frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{4} &< 1 - \frac{1}{2} \cdot \frac{c}{b} \\ \frac{p_{CDC} + p_{DCD}}{2} &< 1 - \frac{1}{2} \cdot \frac{c}{b} \end{aligned} \tag{48}$$

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

$$
\begin{aligned}
0 \quad &\leq \quad s_{\mathbf{m},\mathbf{p}} - (b-c) \\
&\overset{Eq.\ (14)}{=} \quad b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b-c) \\
&\overset{Eqs.\ (47),(34)}{=} \quad b\left(v^2_{CCC}p_{CCC}+v^2_{CCD}(p_{CCD}+p_{DCC})+v^2_{CDC}p_{CDC}+v^2_{DDC}(p_{CDD}+p_{DDC})+v^2_{DCD}p_{DCD}+v^2_{DDD}p_{DDD}\right) \\
&\qquad -c\left(v^2_{CCC}+2v^2_{CCD}+v^2_{DCD}\right) - (b-c)\left(v^2_{CCC}+2v^2_{CCD}+v^2_{CDC}+2v^2_{DDC}+v^2_{DCD}+v^2_{DDD}\right) \\
&= \quad b\, v^2_{CCC}(p_{CCC}-1) + v^2_{CCD}(b\,(p_{CCD}+p_{DCC}-2)) + v^2_{CDC}(b\,(p_{CDC}-1)+c)+ \\
&\qquad v^2_{CDD}(b\,(p_{CDD}+p_{DDC}-2)+2\,c) + v^2_{DCD}(b\,(p_{DCD}-1)) + v^2_{DDD}(b\,(p_{DDD}-1)+c) \\
&\overset{Eq.\ (46)}{=} \quad b\, v^2_{CCC}(p_{CCC}-1) + v^2_{DDD}(b\,(p_{DDD}-1)+c) + v^2_{CDC}(b\,(p_{CCD}+p_{DCC}+p_{CDC}-3)+c)+ \\
&\qquad v^2_{CDD}(b\,(p_{CDD}+p_{DDC}+p_{CCD}+p_{DCC}-4)+2\,c) + v^2_{DCD}(b\,(p_{DCD}-1) - b\,(p_{CCD}+p_{DCC})-2)
\end{aligned}
\tag{49}
$$

Condition (49) 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. (48) reassure that $\mathbf{p}$ is Nash strategy, and given that $p_{CC}=1$, it is a partner strategy. $\square$

C Proof of Corollary 5.2.3

To prove corollary 5.2.3 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,

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

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

Then we define w_i as,

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

Note that,

$$\sum_{i=0}^{r} w_i = 1. \tag{50}$$

The cooperation rate of the reactive player is given by,

$$\rho_{\mathbf{p}} = \sum_{i=0}^{n} r_i \cdot w_i. \tag{51}$$

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,

$$\rho_{\tilde{\mathbf{p}}} = \sum_{i=0}^{n} \frac{i}{n} \cdot w_i. \tag{52}$$

We will also define the self-reactive counting round repeat strategies. These are strategies that start by playing a sequence of cooperation in the first n moves until they reach a total of i number of cooperations. Then they repeat their k moves. We will also define a set of self-reactive $i-$ repeat strategies. That is strategies that repeat sequences where the sequence has a total of i cooperation. These strategies start by playing their sequence and their after repeat their a^{i}_{-n} action. We denote this set of strategies as $A = \{\mathbf{A^0}, \mathbf{A^1}, \dots, \mathbf{A^n}\}$.

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

$$s_{\mathbf{A^i},\mathbf{r}} = b \cdot r_i - \frac{i}{n} \cdot c \quad for \quad i \in [0, n]. \tag{53}$$

The intuition behind Eq. (53) 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 $i \in [0, n]$),

$$s_{\mathbf{A^i},\mathbf{r}} \leq b - c \tag{54}$$

$$b \cdot r_i - \frac{i}{n} \cdot c \leq b - c \tag{55}$$

$$r_i \leq 1 - \frac{i}{n} \cdot \frac{c}{b} \tag{56}$$

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

$$\begin{aligned}
0 \quad & \leq \quad s_{\mathbf{m},\mathbf{p}} - (b-c) \\
& \overset{Eq.\ (14)}{=} \quad b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b-c) \\
& \overset{Eqs.\ (50),(51),(52)}{=} \quad b\sum_{k=0}^{n} r_{n-k} \cdot u_{n-k} - c\sum_{k=0}^{n} \frac{n-k}{n} \cdot u_{n-k} - (b-c)\sum_{k=0}^{n} u_{n-k} \\
& \quad u_n\Big(b\,(r_n - 1)\Big) + \sum_{k=1}^{n} u_{n-k}\Big(b\sum_{k=1}^{n} r_{n-k} - c\sum_{k=0}^{n-1} \frac{n-k}{n} - (b-c)\Big)
\end{aligned} \tag{57}$$

This condition holds only if,

$$\Big(b\,r_{n-k} - c\,\frac{n-k}{n} - (b-c)\Big) < 0 \Rightarrow \tag{58}$$

$$b\,(r_{n-k} - 1) + (1 - \frac{n-k}{n})\,c < 0 \Rightarrow \tag{59}$$

$$r_{n-k} < 1 - \frac{n}{k} \cdot \frac{c}{b}. \tag{60}$$

for $k \in [0, n]$. Thus, any counting strategy that satisfies conditions (54) is Nash, and if it is nice, it's also a partner strategy. $\square$

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.

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.