

Reactive strategies with longer memory

Nikoleta E. Glynnatsi, 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

$$\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}. \quad (1)$$

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,

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix}, \quad (2)$$

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 in the following, 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 $h^i = (a_{-n}^i, \dots, a_{-1}^i) \in \{C, D\}^n$. An entry a_{-k}^i corresponds to player i 's action k rounds ago. Let H^i denote the space of all n -histories for $i \in \{1, 2\}$. 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. This set contains $|H| = 2^{2n}$ elements.

A *memory- n* strategy is a vector $\mathbf{m} = (m_h)_{h \in H} \in [0, 1]^{2^{2n}}$. 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 $1 \leq k \leq n$. 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,

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

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 denoted by 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 during the past n rounds. 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 a memory- n , a reactive- n , and a self-reactive- n strategy.

Let players 1 and 2 use memory- n strategies \mathbf{m} and \mathbf{m}' . Then one can represent the interaction as a Markov chain. The set of states of the chain is the space of all possible histories H . The $2^{2n} \times 2^{2n}$ transition matrix M , describes the transition probabilities between all possible histories. For $h = (h^1, h^2), \tilde{h} = (\tilde{h}^1, \tilde{h}^2) \in H$,

$$M_{h, \tilde{h}} = \begin{cases} m_h \cdot m'_h & \text{if } a_{-k}^1 = C \text{ and } a_{-k}^2 = C \\ m_h \cdot (1 - m'_h) & \text{if } a_{-k}^1 = C \text{ and } a_{-k}^2 = D \\ (1 - m_h) \cdot m'_h & \text{if } a_{-k}^1 = D \text{ and } a_{-k}^2 = C \\ (1 - m_h) \cdot (1 - m'_h) & \text{if } a_{-k}^1 = D \text{ and } a_{-k}^2 = D \\ 0, & \text{if } ((a_{-(n-1)}^1, \dots, a_{-1}^1), (a_{-(n-1)}^2, \dots, a_{-1}^2)) \neq ((\tilde{a}_{-n}^1, \dots, \tilde{a}_{-2}^1), (\tilde{a}_{-n}^2, \dots, \tilde{a}_{-2}^2)) \end{cases}$$

The final case ensures that the only reachable states are those in which the outcomes of these states match the previous outcomes in the 'next' state.

Let $\mathbf{v} = (v_h)_{h \in H}$ be an invariant distribution of this Markov chain. Based on the invariant distribution \mathbf{v} , we can also compute the players' payoffs. To this end, 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,

$$S_h^k = \begin{cases} b - c & \text{if } a_{-k}^1 = C \text{ and } a_{-k}^2 = C \\ -c & \text{if } a_{-k}^1 = C \text{ and } a_{-k}^2 = D \\ b & \text{if } a_{-k}^1 = D \text{ and } a_{-k}^2 = C \\ 0 & \text{if } a_{-k}^1 = D \text{ and } a_{-k}^2 = D \end{cases} \quad (3)$$

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

$$s_{\mathbf{m}, \mathbf{m}'} = \mathbf{v} \cdot \mathbf{S}^1 = \mathbf{v} \cdot \mathbf{S}^2 = \dots = \mathbf{v} \cdot \mathbf{S}^n. \quad (4)$$

The equalities $\mathbf{v} \cdot \mathbf{S}^1 = \dots = \mathbf{v} \cdot \mathbf{S}^n$ correspond to the intuition that it does not matter which of the past n rounds we use to define average payoffs. The payoff $s_{\mathbf{m}', \mathbf{m}}$ of player 2 can be defined analogously.

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,

$$s_{\mathbf{m}', \mathbf{m}} \leq s_{\mathbf{m}, \mathbf{m}} \quad \forall m' \in [0, 1]^{2n}. \quad (5)$$

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 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 $\{\mathbf{v}^k, k = 1, 2, \dots\}$ with \mathbf{v}^k representing the distribution over the states in the k^{th} round of the game. Let \mathbf{v} be an associated stationary distribution of the interaction. In this work Akin shows that,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbf{v}^k \cdot (\mathbf{m} - \mathbf{m}^{1-\text{Rep}}) = 0, \text{ and therefore } \mathbf{v} \cdot (\mathbf{m} - \mathbf{m}^{1-\text{Rep}}) = 0. \quad (6)$$

With the same method as in [Akin, 2016], one can derive a generalised version of his result, Lemma.

Lemma 2.1. Let player 1 use a memory- n strategy, and let player 2 use any arbitrary strategy.

The interaction between the two players leads to.

one can show *Akin's Lemma*: For each k with $1 \leq k \leq n$, the invariant distribution \mathbf{v} satisfies the following relationship,

$$\mathbf{v} \cdot (\mathbf{m} - \mathbf{m}^{k-\text{Rep}}) = \sum_{h \in H} v_h (m_h - m_h^{k-\text{Rep}}) = 0. \quad (7)$$

The intuition for this result is that $\mathbf{v} \cdot \mathbf{m}$ and all $\mathbf{v} \cdot \mathbf{m}^{k-\text{Rep}}$ are just different (but equivalent) expressions for player p 's average cooperation rate. For example, $\mathbf{v} \cdot \mathbf{m}$ corresponds to a setup in which one first draws a history h according to the invariant distribution \mathbf{v} ; then one takes player p 's probability m_h to cooperate in the next round; the expectation of this procedure is $\sum_{h \in H} v_h m_h$.

Proof. The probability that p cooperates in the n^{th} round, denoted by $\hat{v}_C^{(n)}$, is $\hat{v}_C^{(n)} = \hat{v}_1^{(n)} + \hat{v}_2^{(n)} + \hat{v}_5^{(n)} + \hat{v}_6^{(n)} + \hat{v}_9^{(n)} + \hat{v}_{10}^{(n)} + \hat{v}_{13}^{(n)} + \hat{v}_{14}^{(n)} = \hat{\mathbf{v}} \cdot \hat{\mathbf{e}}_{12}$ where $\hat{\mathbf{e}}_{12} = (1, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0)$. The probability that p cooperates in the $(n+1)^{\text{th}}$ round, denoted by $\hat{v}_C^{(n+1)} = \hat{\mathbf{v}}^{(n)} \cdot \hat{\mathbf{p}}$. Hence,

$$\hat{v}_C^{(n+1)} - \hat{v}_C^{(n)} = \hat{\mathbf{v}}^{(n)} \cdot \hat{\mathbf{p}} - \hat{\mathbf{v}} \cdot \hat{\mathbf{e}}_{12} = \hat{\mathbf{v}}^{(n)} \cdot (\hat{\mathbf{p}} - \hat{\mathbf{e}}_{12}) = \hat{\mathbf{v}}^{(n)} \cdot \tilde{\mathbf{p}}.$$

This implies,

$$\sum_{k=1}^n \hat{\mathbf{v}}^{(k)} \cdot \tilde{\mathbf{p}} = \sum_{k=1}^n \hat{v}_C^{(k+1)} - \hat{v}_C^{(k)} \Rightarrow \sum_{k=1}^n \hat{\mathbf{v}}^{(k)} \cdot \tilde{\mathbf{p}} = \hat{v}_C^{(n+1)} - \hat{v}_C^{(1)}. \quad (8)$$

As the right side has absolute value at most 1,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \hat{\mathbf{v}}^{(k)} \cdot \tilde{\mathbf{p}} = 0. \quad (9)$$

□

Zero-determinant strategies. Based on Akin's Lemma, 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} is a zero-determinant strategy if there are k_1, k_2, k_3 and α, β, γ such that \mathbf{m} can be written as

$$\mathbf{m} = \alpha \mathbf{S}^{k_1} + \beta \tilde{\mathbf{S}}^{k_2} + \gamma \mathbf{1} + \mathbf{m}^{k-\text{Rep}}, \quad (10)$$

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

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

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

$$m_h = \begin{cases} 1 & \text{if } a_{-k}^q = C \\ 0 & \text{if } a_{-k}^q = D \end{cases} \quad (12)$$

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 $s_{\mathbf{m}, \mathbf{m}'} = s_{\mathbf{m}', \mathbf{m}}$.

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

$$m_h = \begin{cases} 1 & \text{if } a_{-k}^q = C \\ 1 - c/b & \text{if } a_{-k}^q = D \end{cases} \quad (13)$$

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

3 Reactive Partner Strategies

A *reactive- n strategy* is denoted by a vector $\mathbf{p} = (p_h)_{h \in H^q} \in [0, 1]^{2^n}$. Each entry p_h corresponds to the player's cooperation probability in the next round, based on the co-player's action(s) in the previous n rounds. Therefore, n -bit reactive strategies exclusively rely on the co-player's n -history, remaining unaffected by the focal player's own actions during the past n rounds. From this point onward, we distinguish between memory- n strategies and reactive- n strategies, using notations \mathbf{m} and \mathbf{p} respectively for each set of strategies.

To begin, let's introduce some additional notation. Suppose player p adopts a reactive- n strategy \mathbf{p} , and suppose player q adopts an arbitrary memory- n strategy. Let $\mathbf{v} = (v_h)_{h \in H}$ be an invariant of the game between the two players. We define the following marginal distributions with respect to the possible n -histories of player q ,

$$v_h^q = \sum_{h^p \in H^p} v_{(h^p, h^q)} \quad \forall h^q \in H^q. \quad (14)$$

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

$$\sum_{h \in H^q} v_h^q = 1. \quad (15)$$

Similarly, the cooperation rate of player q can also be defined irrespective of the actions of player p . Let H_k^q be the subset of H^q , for which,

$$H_k^q = \{h \in H^q : h_{-k} = C\}. \quad (16)$$

Let $\rho_{\mathbf{m}}$ be the cooperation rate of player q playing an arbitrary memory- n strategy \mathbf{m} when playing against player p with a reactive strategy,

$$\rho_{\mathbf{m}} = \sum_{h \in H_1^q} v_h^q = \sum_{h \in H_2^q} v_h^q = \dots = \sum_{h \in H_n^q} v_h^q. \quad (17)$$

Equality (17) corresponds to the intuition that it does not matter which of the past n rounds player q cooperated to define the cooperation rate.

We can also express player p 's average cooperation rate $\rho_{\mathbf{p}}$ in terms of v_h^q by noting that,

$$\rho_{\mathbf{p}} = \sum_{h \in H^q} v_h^q \cdot p_h. \quad (18)$$

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,

$$\begin{aligned} s_{\mathbf{p}, \mathbf{m}} &= b \rho_{\mathbf{m}} - c \rho_{\mathbf{p}} \\ s_{\mathbf{m}, \mathbf{q}} &= b \rho_{\mathbf{p}} - c \rho_{\mathbf{m}}. \end{aligned} \quad (19)$$

3.1 Sufficiency of Self reactive strategies

To characterize all partner reactive- n strategies, one would usually need to check against all pure memory- n strategies McAvooy and Nowak [2019]. However, we demonstrate that when player p uses a reactive- n strategy, it is sufficient to check only against self-reactive- n strategies. This is a direct outcome of Lemma 3.1.

Self-reactive- n strategies are also a subset of memory- n strategies. They only consider the focal player's own n -history, and ignore the co-player's. Formally, a self-reactive- n strategy is a vector $\tilde{\mathbf{p}} = (\tilde{p}_h)_{h \in H^p} \in [0, 1]^{2^n}$. Each entry \tilde{p}_h corresponds to the player's cooperation probability in the next, depending on the player's own action(s) in the previous n rounds.

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

Proof. □

Note that Lemma 3.1 aligns with the previous result by Press and Dyson [2012]. They 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. The result here 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 short-memory player.

More specifically, the play of a self-reactive player solely relies on their own previous actions. Hence, describing the self-reactive player's play can be achieved through a Markov process with a $2^n \times 2^n$ transition matrix \tilde{M} instead. The stationary distribution $\tilde{\mathbf{v}}$ of \tilde{M} has the following property:

$$\tilde{u}_h = u_h^q \forall h \in H^q. \quad (20)$$

From hereupon we will use the notation \mathbf{m}, \mathbf{p} , and $\tilde{\mathbf{p}}$ to denote memory- n , reactive- n , and self-reactive- n strategies.

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 3.2 (“Reactive-Two Partner Strategies”). A reactive-two strategy \mathbf{p} , is a partner strategy if and only if, it's nice ($p_{CC} = 1$) and the remaining entries satisfy the conditions:

$$p_{DD} < 1 - \frac{c}{b} \quad \text{and} \quad \frac{p_{CD} + p_{DC}}{2} < 1 - \frac{1}{2} \cdot \frac{c}{b}. \quad (21)$$

There are two independent proves of Theorem 3.2. The first prove is in line with the work of [Akin, 2016], and the second one relies on Lemma 3.1. Here, we discuss both.

Proof One. Suppose player p adopts a reactive-two strategy $\mathbf{p} = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$. Moreover, suppose player q 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.

We define the following four marginal distributions with respect to the possible two-histories of player q ,

$$\begin{aligned} v_{CC}^q &= \sum_{h^p \in H^p} v_{(h^p, CC)} \\ v_{CD}^q &= \sum_{h^p \in H^p} v_{(h^p, CD)} \\ v_{DC}^q &= \sum_{h^p \in H^p} v_{(h^p, DC)} \\ v_{DD}^q &= \sum_{h^p \in H^p} v_{(h^p, DD)}. \end{aligned} \tag{22}$$

These four entries describe how often we observe player q to choose actions CC , CD , DC , DD in two consecutive rounds (irrespective of the actions of player p). We can define player q 's average cooperation rate $\rho_{\mathbf{m}}$ as

$$\rho_{\mathbf{m}} := v_{CC}^q + v_{CD}^q = v_{CC}^q + v_{DC}^q. \tag{23}$$

Here, the second equality holds because it does not matter whether we define player q 's cooperation rate based on the first or the second round of each 2-history. In particular, we can use this equality to conclude

$$v_{CD}^q = v_{DC}^q. \tag{24}$$

Similarly, we can express player p 's average cooperation rate $\rho_{\mathbf{p}}$ in terms of v_{CC}^q , v_{CD}^q , v_{DC}^q , v_{DD}^q by noting that

$$\begin{aligned} \rho_{\mathbf{p}} &= v_{CC}^q p_{CC} + v_{CD}^q p_{CD} + v_{DC}^q p_{DC} + v_{DD}^q p_{DD} \\ &= v_{CC}^q p_{CC} + v_{CD}^q (p_{CD} + p_{DC}) + v_{DD}^q p_{DD}. \end{aligned} \tag{25}$$

Here, the second equality is due to Eq. (24).

Proof. 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 AllD = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0) 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

$$p_{DD} \leq 1 - \frac{c}{b} \quad \text{and} \quad p_{CD} + p_{DC} \leq 1 + \frac{b - c}{c}. \tag{26}$$

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

$$\begin{aligned}
0 &< s_{\mathbf{m},\mathbf{p}} - (b-c) \\
&\stackrel{\text{Eq. (19)}}{=} b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b-c) \\
&\stackrel{\text{Eqs. (23),(25),(15)}}{=} b \left(v_{CC}^q p_{CC} + v_{CD}^q (p_{CD} + p_{DC}) + v_{DD}^q p_{DD} \right) - c \left(v_{CC}^q + v_{CD}^q \right) - (b-c) \left(v_{CC}^q + 2v_{CD}^q + v_{DD}^q \right) \\
&= v_{CC}^q b (p_{CC} - 1) + v_{CD}^q \left(b(p_{CD} + p_{DC}) + c - 2b \right) + v_{DD}^q \left(bp_{DD} - (b-c) \right).
\end{aligned} \tag{27}$$

Condition (27) can hold only if,

$$b(p_{CD} + p_{DC}) + c - 2b > 0, \quad bp_{DD} - (b-c) > 0. \tag{28}$$

Thus, Eq. (26) reassures that \mathbf{p} is Nash strategy, and given that $p_{CC} = 1$, it is a partner strategy.

□

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

$$s_{\mathbf{m},\mathbf{p}} \leq (b-c), \tag{29}$$

must hold against all pure memory-2 strategies ($\mathbf{m} \in \{0, 1\}^{4^2}$). Due to Lemma 3.1, it is sufficient to check only against pure self-reactive strategies, and in the case of $n = 2$ there can be only 16 such strategies. We refer to them as $\tilde{\mathbf{q}}^i$ for $i \in 1, \dots, 16$. The strategies are as follow,

• $\tilde{\mathbf{q}}^0 = (0, 0, 0, 0)$	• $\tilde{\mathbf{q}}^4 = (0, 1, 0, 0)$	• $\tilde{\mathbf{q}}^8 = (1, 0, 0, 0)$	• $\tilde{\mathbf{q}}^{12} = (1, 1, 0, 0)$
• $\tilde{\mathbf{q}}^1 = (0, 0, 0, 1)$	• $\tilde{\mathbf{q}}^5 = (0, 1, 0, 1)$	• $\tilde{\mathbf{q}}^9 = (1, 0, 0, 1)$	• $\tilde{\mathbf{q}}^{13} = (1, 1, 0, 1)$
• $\tilde{\mathbf{q}}^2 = (0, 0, 1, 0)$	• $\tilde{\mathbf{q}}^6 = (0, 1, 1, 0)$	• $\tilde{\mathbf{q}}^{10} = (1, 0, 1, 0)$	• $\tilde{\mathbf{q}}^{14} = (1, 1, 1, 0)$
• $\tilde{\mathbf{q}}^3 = (0, 0, 1, 1)$	• $\tilde{\mathbf{q}}^7 = (0, 1, 1, 1)$	• $\tilde{\mathbf{q}}^{11} = (1, 0, 1, 1)$	• $\tilde{\mathbf{q}}^{15} = (1, 1, 1, 1)$

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

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

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

$$p_{DD} \leq 1 - \frac{c}{b} \quad (30)$$

$$\frac{p_{CD} + p_{DC} + p_{DD}}{3} \leq 1 - \frac{2}{3} \cdot \frac{c}{b} \quad (31)$$

$$\frac{p_{CD} + p_{DC}}{2} \leq 1 - \frac{1}{2} \cdot \frac{c}{b} \quad (32)$$

$$(33)$$

Note that condition (32) is the sum of conditions (31) and (33). Thus, only conditions (31) and (33) are necessary. □

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 is represented by the vector $\mathbf{p} = (1, p_{CCD}, p_{CDC}, p_{CDD}, p_{DCC}, p_{DCD}, p_{DDC}, p_{DDD})$.

Theorem 3.3 (“Reactive-Three Partner Strategies”). A reactive-three strategy \mathbf{p} , is a partner strategy if and only if, it's nice ($p_{CCC} = 1$) and the remaining entries satisfy the conditions:

$$\frac{p_{CCD} + p_{CDC} + p_{DCC}}{3} < 1 - \frac{1}{3} \cdot \frac{c}{b} \quad \frac{p_{CDD} + p_{DCD} + p_{DDC}}{3} < 1 - \frac{2}{3} \cdot \frac{c}{b} \quad p_{DDD} < 1 - \frac{c}{b} \quad (34)$$

$$\frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{4} < 1 - \frac{1}{2} \cdot \frac{c}{b} \quad \frac{p_{CDC} + p_{DCD}}{2} < 1 - \frac{1}{2} \cdot \frac{c}{b} \quad (35)$$

Once again, there are two independent proves of Theorem 3.3, and we will present both.

Proof One. Suppose player p adopts a reactive-three strategy \mathbf{p} , and suppose player q 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.

We define the following eight marginal distributions with respect to the possible three-histories of player q ,

$$\begin{aligned}
v_{CCC}^q &= \sum_{h^p \in H^p} v_{(h^p, CCC)} \\
v_{CCD}^q &= \sum_{h^p \in H^p} v_{(h^p, CCD)} \\
v_{CDC}^q &= \sum_{h^p \in H^p} v_{(h^p, CDC)} \\
v_{CDD}^q &= \sum_{h^p \in H^p} v_{(h^p, CDD)} \\
v_{DCC}^q &= \sum_{h^p \in H^p} v_{(h^p, DCC)} \\
v_{DCD}^q &= \sum_{h^p \in H^p} v_{(h^p, DCD)} \\
v_{DDC}^q &= \sum_{h^p \in H^p} v_{(h^p, DDC)} \\
v_{DDD}^q &= \sum_{h^p \in H^p} v_{(h^p, DDD)}.
\end{aligned} \tag{36}$$

These eight entries describe how often we observe player q to choose actions CCC , CCD , CDC , CDD , DCC , DCD , DDC , DDD in three consecutive rounds (irrespective of the actions of player p). We can define player q 's average cooperation rate $\rho_{\mathbf{m}}$ as

$$\rho_{\mathbf{m}} := v_{CCC}^q + v_{CCD}^q + v_{DCC}^q + v_{DCD}^q. \tag{37}$$

In the case of $n = 3$ the following equalities hold,

$$v_{CCD}^q = v_{DCC}^q \tag{38}$$

$$v_{DDC}^q = v_{CDD}^q \tag{39}$$

$$\begin{aligned}
v_{CCD}^q + v_{DCD}^q &= v_{CDC}^q + v_{DDC}^q \Rightarrow \\
v_{CCD}^q &= v_{CDC}^q + v_{CDD}^q - v_{DCD}^q
\end{aligned} \tag{40}$$

The average cooperation rate of p 's is given by

$$\begin{aligned}
\rho_{\mathbf{p}} &= v_{CCC}^q p_{CCC} + v_{CCD}^q p_{CCD} + v_{CDC}^q p_{CDC} + v_{CDD}^q p_{CDD} + v_{DCC}^q p_{DCD} + \\
&\quad + v_{DDC}^q p_{DDC} + v_{DDD}^q p_{DDD} \\
&\stackrel{\text{Eqs. (38), (39)}}{=} v_{CCC}^q p_{CCC} + v_{CCD}^q (p_{CCD} + p_{DCC}) + v_{CDC}^q p_{CDC} + v_{CDD}^q (p_{CDD} + p_{DDC}) + \\
&\quad + v_{DCC}^q p_{DCD} + v_{DDD}^q p_{DDD}
\end{aligned} \tag{41}$$

Proof. 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, $\tilde{\mathbf{q}}^{15} = (0, 0, 0, 0, 1, 1, 1, 1)$, $\tilde{\mathbf{q}}^{17} = (0, 0, 0, 1, 0, 0, 0, 1)$, $\tilde{\mathbf{q}}^{51} = (0, 0, 1, 1, 0, 0, 1, 1)$ and $\tilde{\mathbf{q}}^{102} = (0, 1, 1, 0, 0, 1, 1, 0)$. Thus, \mathbf{p} can only form a Nash equilibrium if

$$\pi(\text{AllD}, \mathbf{p}) \leq b-c \quad \text{and} \quad \pi(\tilde{\mathbf{q}}^{15}, \mathbf{p}) \leq b-c \quad \text{and} \quad \pi(\tilde{\mathbf{q}}^{17}, \mathbf{p}) \leq b-c \quad \text{and}$$

$$\pi(\tilde{\mathbf{q}}^{51}, \mathbf{p}) \leq b-c \quad \text{and} \quad \pi(\tilde{\mathbf{q}}^{102}, \mathbf{p}) \leq b-c$$

or equivalently, if

$$\frac{p_{CCD} + p_{CDC} + p_{DCC}}{3} < 1 - \frac{1}{3} \cdot \frac{c}{b} \quad \frac{p_{CDD} + p_{DCD} + p_{DDC}}{3} < 1 - \frac{2}{3} \cdot \frac{c}{b} \quad p_{DDD} < 1 - \frac{c}{b} \quad (42)$$

$$\frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{4} < 1 - \frac{1}{2} \cdot \frac{c}{b} \quad \frac{p_{CDC} + p_{DCD}}{2} < 1 - \frac{1}{2} \cdot \frac{c}{b} \quad (43)$$

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

$$\begin{aligned} 0 &\leq s_{\mathbf{m}, \mathbf{p}} - (b-c) \\ &\stackrel{\text{Eq. (19)}}{=} b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b-c) \\ &\stackrel{\text{Eqs. (41), (15)}}{=} b \left(v_{CCC}^q p_{CCC} + v_{CCD}^q (p_{CCD} + p_{DCC}) + v_{CDC}^q p_{CDC} + v_{DDC}^q (p_{CDD} + p_{DDC}) + v_{DCD}^q p_{DCD} + v_{DDD}^q p_{DDD} \right) \\ &\quad - c \left(v_{CCC}^q + 2v_{CCD}^q + v_{DCD}^q \right) - (b-c) \left(v_{CCC}^q + 2v_{CCD}^q + v_{CDC}^q + 2v_{DDC}^q + v_{DCD}^q + v_{DDD}^q \right) \\ &= b v_{CCC}^q (p_{CCC} - 1) + v_{CCD}^q (b(p_{CCD} + p_{DCC} - 2)) + v_{CDC}^q (b(p_{CDC} - 1) + c) + \\ &\quad v_{DCD}^q (b(p_{CDD} + p_{DDC} - 2) + 2c) + v_{DDC}^q (b(p_{DDC} - 1)) + v_{DDD}^q (b(p_{DDD} - 1) + c) \\ &\stackrel{\text{Eq. (40)}}{=} b v_{CCC}^q (p_{CCC} - 1) + v_{DDD}^q (b(p_{DDD} - 1) + c) + v_{CDC}^q (b(p_{CCD} + p_{DCC} + p_{CDC} - 3) + c) + \\ &\quad v_{DCD}^q (b(p_{CDD} + p_{DDC} + p_{CCD} + p_{DCC} - 4) + 2c) + v_{DDC}^q (b(p_{DDC} - 1) - b(p_{CCD} + p_{DCC}) - 2) \\ &\quad (44) \end{aligned}$$

Condition (44) holds only for,

$$\begin{aligned} &b(p_{DDD} - 1) + c < 0, \quad b(p_{CCD} + p_{DCC} + p_{CDC} - 3) + c \\ &b(p_{CDD} + p_{DDC} + p_{CCD} + p_{DCC} - 4) + 2c < 0 \Rightarrow -b(p_{CCD} + p_{DCC} - 2) > b(p_{CDD} + p_{DDC} - 2) + 2c \\ &b(p_{DCD} - 1) - b(p_{CCD} + p_{DCC}) - 2 < 0 \Rightarrow b(p_{DCD} + p_{CDD} + p_{DDC} - 3) + 2c < 0. \end{aligned}$$

Thus, conditions Eq. (42) reassure that \mathbf{p} is Nash strategy, and given that $p_{CC} = 1$, it is a partner strategy. \square

Proof Two. Consider all the pure self-reactive-three strategies. There is a total of 256 such strategies. These are given in the Section 5. 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 subsection.

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

$$\begin{aligned}
s_{\bar{q}^i, p} &= b p_{DDD} & \text{for } i \in [0, 2, 4, 6, \dots, 250, 252, 254] \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CDD} + p_{DCD} + p_{DDC} + p_{DDD})}{4} - \frac{1}{4} \cdot c & \text{for } i \in \{1, 9, 33, 41, 65, 73, 97, 105, 129, 137, 161, 169, 193, 201, 225, 233\} \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC} + p_{DDD})}{5} - \frac{2}{5} \cdot c & \text{for } i \in \{3, 7, 35, 39, 131, 135, 163, 167\} \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CDC} + p_{DCD})}{2} - \frac{1}{2} \cdot c & \text{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_{\bar{q}^i, p} &= \frac{b \cdot (p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC} + p_{DDD} + 1)}{6} - \frac{1}{2} \cdot c & \text{for } i \in \{11, 15, 43, 47\} \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CDD} + p_{DCD} + p_{DDC})}{3} - \frac{1}{3} \cdot c & \text{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_{\bar{q}^i, p} &= \frac{b \cdot (p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC})}{4} - \frac{1}{2} \cdot c & \text{for } i \in \{18, 19, 22, 23, 50, 51, 54, 55, 146, 147, 150, 151, 178, 179, 182, 183\} \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC} + 1)}{5} - \frac{3}{5} \cdot c & \text{for } i \in \{26, 27, 30, 31, 58, 59, 62, 63\} \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC} + p_{DDD})}{7} - \frac{3}{7} \cdot c & \text{for } i \in \{37, 67, 165, 195\} \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC} + p_{DDD} + 1)}{8} - \frac{1}{2} \cdot c & \text{for } i \in \{45, 75\} \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC})}{6} - \frac{1}{2} \cdot c & \text{for } i \in \{52, 53, 82, 83, 180, 181, 210, 211\} \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC} + 1)}{7} - \frac{4}{7} \cdot c & \text{for } i \in \{60, 61, 90, 91\} \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CCD} + p_{CDC} + p_{DCC})}{3} - \frac{2}{3} \cdot c & \text{for } i \in \{96-103, 112-119, 224-231, 240-247\} \\
s_{\bar{q}^i, p} &= \frac{b \cdot (p_{CCD} + p_{CDC} + p_{DCC} + 1)}{4} - \frac{3}{4} \cdot c & \text{for } i \in \{104-111, 120-127\} \\
s_{\bar{q}^i, p} &= (b - c) & \text{for } i \in [128, 255]
\end{aligned} \tag{45}$$

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

$$p_{DDD} \leq 1 - \frac{c}{b}, \quad \frac{p_{CDC} + p_{DCD}}{2} \leq 1 - \frac{1}{2} \cdot \frac{c}{b}, \quad \frac{p_{CDD} + p_{DCD} + p_{DDC}}{3} \leq 1 - \frac{2}{3} \cdot \frac{c}{b}, \tag{46}$$

$$\frac{p_{CCD} + p_{CDC} + p_{DCC}}{3} \leq 1 - \frac{1}{3} \cdot \frac{c}{b}, \quad \frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{4} \leq 1 - \frac{1}{2} \cdot \frac{c}{b} \tag{47}$$

$$\frac{p_{CDD} + p_{DCD} + p_{DDC} + p_{DDD}}{4} \leq 1 - \frac{3}{4} \cdot \frac{c}{b}, \tag{48}$$

$$\frac{p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC} + p_{DDD}}{7} \leq 1 - \frac{4}{7} \cdot \frac{c}{b}, \tag{49}$$

$$\frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC} + p_{DDD}}{5} \leq 1 - \frac{3}{5} \cdot \frac{c}{b}, \tag{50}$$

$$\frac{p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC}}{6} \leq 1 - \frac{1}{2} \cdot \frac{c}{b} \tag{51}$$

Note that only conditions (46) and (47) are unique. The remaining conditions can be derived from the sums of two or more of these conditions. \square

3.4 Reactive Counting Partner 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-counting- n strategy is represented by a vector $\mathbf{r} = (r_i)_{i \in [0, n]}$, where the entry r_i indicates the probability of cooperating given that the co-player cooperated i times in the last n turns.

Reactive-Counting-Two Partner Strategies. These are denoted by the vector $\mathbf{r} = (r_2, r_1, r_0)$ where r_i is the probability of cooperating in after i cooperations in the last 2 turns. We can characterise reactive-counting-two partner strategies by setting $r_2 = 1$, and $p_{CD} = p_{DC} = r_1$ and $p_{DD} = r_0$ in conditions (21). This gives us the following result.

Lemma 3.4. A nice reactive-counting-two strategy $\mathbf{r} = (1, r_1, r_0)$ is a partner strategy if and only if,

$$r_1 < 1 - \frac{1}{2} \cdot \frac{c}{b} \text{ and } r_0 < 1 - \frac{c}{b}. \quad (52)$$

Reactive-Counting-Three Partner Strategies. These are denoted by the vector $\mathbf{r} = (r_3, r_2, r_1, r_0)$ where r_i is the probability of cooperating in after i cooperations in the last 3 turns. We can characterise reactive-counting-three partner strategies by setting $r_3 = 1$, and $p_{CCD} = p_{CDC} = r_2, p_{DCD} = p_{DDC} = r_1$ and $p_{DDD} = r_0$ in conditions (34). This gives us the following result.

Lemma 3.5. A nice reactive-counting-three strategy $\mathbf{r} = (1, r_2, r_1, r_0)$ is a partner strategy if and only if,

$$r_2 < 1 - \frac{1}{3} \cdot \frac{c}{b}, \quad r_1 < 1 - \frac{2}{3} \cdot \frac{c}{b} \text{ and } r_0 < 1 - \frac{c}{b}. \quad (53)$$

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

Given a reactive-counting- n strategy $\mathbf{r} = (r_n, r_{n-1}, \dots, r_0)$, in the strategy's eyes the game can end up in n states. Each u_i state represents the state that the co-player cooperated i times in the last n turns with,

$$\sum_{i=0}^n u_i = 1 \Rightarrow \sum_{k=0}^n u_{n-k} = 1. \quad (54)$$

Thus the cooperation ratio of the strategy is,

$$\rho_{\mathbf{P}} = \sum_{k=0}^n r_{n-k} \cdot u_{n-k}. \quad (55)$$

the probability of cooperating given that the co-player cooperated i times. 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 u_i . More specifically, the co-player's cooperation rate is,

$$\rho_{\tilde{\mathbf{P}}} = \sum_{k=0}^n \frac{n-k}{n} \cdot u_{n-k}. \quad (56)$$

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

Theorem 3.6 (“Reactive-Counting Partner Strategies”). A reactive-counting- n strategy $\mathbf{r} = (r_i)_{i \in [0, n]}$, is a partner strategy if and only if, the r_i entries satisfy the conditions:

$$r_n = 1, \text{ and } r_{n-k} < 1 - \frac{k}{n} \cdot \frac{c}{b}, \text{ for } k \in [1, n]. \quad (57)$$

Proof. Consider a set of alternating self-reactive- n strategies that defect after cooperating i times. Since $i \in [0, n]$, there can be only $n + 1$ such strategies. We will denote this set 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,

$$s_{\mathbf{A}^i, \mathbf{r}} = b \cdot r_i - \frac{i}{n} \cdot c \text{ for } i \in [0, n]. \quad (58)$$

The intuition behind Eq. (58) is that in the long term of the game the strategies end up in a state where \mathbf{A}^i has cooperated i times in the last n turns. Thus, here the co-player will cooperate, and provide the benefit b with a probability r_i . Whilst in return the alternating strategy cooperated $\frac{i}{n}$ and pays the cost. As we have already discussed previously, a strategy can only be Nash if the payoff of the co-player does not exceed $(b - c)$. Thus for \mathbf{p} to be Nash against each strategy in set A (for $i \in [0, n]$),

The intuition behind Eq. (58) 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. 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 \mathbf{p} to be a Nash equilibrium against each strategy in set A (for $i \in [0, n]$),

$$s_{\mathbf{A}^i, \mathbf{r}} \leq b - c \quad (59)$$

$$b \cdot r_i - \frac{i}{n} \cdot c \leq b - c \quad (60)$$

$$r_i \leq 1 - \frac{i}{n} \cdot \frac{c}{b} \quad (61)$$

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

$$\begin{aligned} 0 &\leq s_{\mathbf{m}, \mathbf{p}} - (b - c) \\ &\stackrel{\text{Eq. (19)}}{=} b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b - c) \\ &\stackrel{\text{Eqs. (54), (55), (56)}}{=} b \sum_{k=0}^n r_{n-k} \cdot u_{n-k} - c \sum_{k=0}^n \frac{n-k}{n} \cdot u_{n-k} - (b - c) \sum_{k=0}^n u_{n-k} \\ &\quad u_n \left(b(r_n - 1) \right) + \sum_{k=1}^n u_{n-k} \left(b \sum_{k=1}^n r_{n-k} - c \sum_{k=0}^{n-1} \frac{n-k}{n} - (b - c) \sum_{k=0}^{n-1} 1 \right) \end{aligned} \quad (62)$$

This condition holds only if,

$$\left(b r_{n-k} - c \frac{n-k}{n} - (b-c)\right) < 0 \Rightarrow \quad (63)$$

$$b(r_{n-k} - 1) + \left(1 - \frac{n-k}{n}\right)c < 0 \Rightarrow \quad (64)$$

$$r_{n-k} < 1 - \frac{n}{k} \cdot \frac{c}{b}. \quad (65)$$

for $k \in [0, n]$. Thus, any counting strategy that satisfies conditions (59) is Nash, and if it is nice, it's also a partner strategy. \square

4 Prisoner's Dilemma

To characterise partner strategies for the general prisoner's dilemma, we can use the method based on Lemma 3.1. Here we discuss this result in the case of $n = 2$.

There are 16 pure-self reactive strategies in $n = 2$. To calculate the explicit payoff expressions for each pure strategy against a nice reactive-two strategy $\mathbf{p} = (1, p_{CD}, p_{DC}, p_{DD})$ we use the method discussed in Section 3.1. More specifically, for a self-reactive strategy \mathbf{q} , we calculate where the strategy is in the long term using the transition matrix,

$$\tilde{M} = \begin{bmatrix} \tilde{p}_1 & 1 - \tilde{p}_1 & 0 & 0 \\ 0 & 0 & \tilde{p}_2 & 1 - \tilde{p}_2 \\ \tilde{p}_3 & 1 - \tilde{p}_3 & 0 & 0 \\ 0 & 0 & \tilde{p}_4 & 1 - \tilde{p}_4 \end{bmatrix} \quad (66)$$

Using the stationary vector $\tilde{\mathbf{v}}$ we can define the payoffs in the general prisoner's dilemma as follows:

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

$$\begin{aligned} a_R &= \tilde{v}_{CC} p_{CC} \tilde{q}_{CC} + \tilde{v}_{CD} p_{CD} \tilde{q}_{CD} + \tilde{v}_{DC} p_{DC} \tilde{q}_{DC} + \tilde{v}_{DD} p_{DD} \tilde{q}_{DD}, \\ a_S &= \tilde{v}_{CC} p_{CC} (1 - \tilde{q}_{CC}) + \tilde{v}_{CD} p_{CD} (1 - \tilde{q}_{CD}) + \tilde{v}_{DC} p_{DC} (1 - \tilde{q}_{DC}) + \tilde{v}_{DD} p_{DD} (1 - \tilde{q}_{DD}), \\ a_T &= \tilde{v}_{CC} (1 - p_{CC}) \tilde{q}_{CC} + \tilde{v}_{CD} (1 - p_{CD}) \tilde{q}_{CD} + \tilde{v}_{DC} (1 - p_{DC}) \tilde{q}_{DC} + \tilde{v}_{DD} (1 - p_{DD}) \tilde{q}_{DD}, \\ a_P &= \tilde{v}_{CC} (1 - p_{CC}) (1 - \tilde{q}_{CC}) + \tilde{v}_{CD} (1 - p_{CD}) (1 - \tilde{q}_{CD}) + \tilde{v}_{DC} (1 - p_{DC}) (1 - \tilde{q}_{DC}) + \tilde{v}_{DD} (1 - p_{DD}) (1 - \tilde{q}_{DD}). \end{aligned}$$

This gives the following payoff expressions:

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

Setting the above expressions to smaller than R gives the following conditions,

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

Consider the case where $T = 1$ and $S = 0$,

$$\begin{aligned}
(P - 1)p_{DD} &< P - R, \quad (1 - P)(p_{CD} + p_{DC}) + Rp_{DD} < 3R - 2P, \\
(1 - P)p_{CD} + R(p_{CD} + p_{DC}) &< 4R - P - 1, \quad Rp_{CD} + (1 - P)p_{DC} < 2R - P \\
R(p_{CD} + p_{DC}) &< 3R - 1
\end{aligned}$$

$$\begin{aligned}
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CDD}+p_{DCD}+p_{DDC})+3P+(R-S)p_{DDD}+S}{4} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)p_{CDC}+P+(R-S)p_{DCD}+S}{2} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= -P(p_{DDD}-1)+Tp_{DDD} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CCD}+p_{CDD}+p_{DDC})+3P+(R-S)(p_{CDC}+p_{DCC}+p_{DCD}+p_{DDD})+4S+T}{8} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)p_{DCC}+P+(R-S)(p_{CDC}+p_{CCD})+2S}{3} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CCD}+p_{DCC}+p_{DDC})+3P+(R-S)(p_{CDC}+p_{CDD}+p_{DCD})+3S}{6} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CCD}+p_{DDC})+2P+T+(R-S)(p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD})+4S}{7} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CCD}+p_{CDD}+p_{DCC})+3P+(R-S)(p_{DDC}+p_{DDD})+2S}{5} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{DCD}+p_{DDC})+2P+(R-S)p_{CDD}+S}{3} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= R & \text{for } i \in [128-255] \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)p_{CCD}+P+T+(R-S)(p_{CDD}+p_{DCC}+p_{DDC})+S}{5} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CCD}+p_{DCC})+2P+(R-S)(p_{CDD}+p_{DDD})+2S}{4} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)p_{CDC}+2P+T+(R-S)(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC})+4S}{7} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CDC}+p_{CDD}+p_{DCD})+3P+T+(R-S)(p_{CCD}+p_{DCC}+p_{DDC}+p_{DDD})+4S}{8} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CDC}+p_{DDC}+p_{DCD})+3P+T+(R-S)(p_{CCD}+p_{CDD}+p_{DDC})+3S}{6} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC})+4P+(R-S)(p_{CDC}+p_{DCD}+p_{DDD})+3S}{7} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{T+(R-S)(p_{CCD}+p_{CDC}+p_{DCC})+3S}{4} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CCD}+p_{CDD})+2P+T+(R-S)(p_{DCC}+p_{DDC}+p_{DDD})+3S}{6} & \text{for } i \in \\
s_{\tilde{q}^i, \mathbf{p}} &= \frac{(T-P)(p_{CDC}+p_{CDD}+p_{DCC}+p_{DCD})+4P+(R-S)(p_{CCD}+p_{DDC}+p_{DDD})+3S}{7} & \text{for } i \in
\end{aligned}$$

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

5 Pure Self-Reactive-Three Strategies

The 256 pure self-reactive-three strategies and their vectors are as follows,

- $\tilde{\mathbf{q}}^0 = (0, 0, 0, 0, 0, 0, 0, 0)$
- $\tilde{\mathbf{q}}^1 = (0, 0, 0, 0, 0, 0, 0, 1)$
- $\tilde{\mathbf{q}}^2 = (0, 0, 0, 0, 0, 0, 1, 0)$
- $\tilde{\mathbf{q}}^3 = (0, 0, 0, 0, 0, 0, 1, 1)$
- $\tilde{\mathbf{q}}^4 = (0, 0, 0, 0, 0, 1, 0, 0)$
- $\tilde{\mathbf{q}}^5 = (0, 0, 0, 0, 0, 1, 0, 1)$
- $\tilde{\mathbf{q}}^6 = (0, 0, 0, 0, 0, 1, 1, 0)$
- $\tilde{\mathbf{q}}^7 = (0, 0, 0, 0, 0, 1, 1, 1)$
- $\tilde{\mathbf{q}}^8 = (0, 0, 0, 0, 1, 0, 0, 0)$
- $\tilde{\mathbf{q}}^9 = (0, 0, 0, 0, 1, 0, 0, 1)$
- $\tilde{\mathbf{q}}^{10} = (0, 0, 0, 0, 1, 0, 1, 0)$
- $\tilde{\mathbf{q}}^{11} = (0, 0, 0, 0, 1, 0, 1, 1)$
- $\tilde{\mathbf{q}}^{12} = (0, 0, 0, 0, 1, 1, 0, 0)$
- $\tilde{\mathbf{q}}^{13} = (0, 0, 0, 0, 1, 1, 0, 1)$
- $\tilde{\mathbf{q}}^{14} = (0, 0, 0, 0, 1, 1, 1, 0)$
- $\tilde{\mathbf{q}}^{15} = (0, 0, 0, 0, 1, 1, 1, 1)$
- $\tilde{\mathbf{q}}^{16} = (0, 0, 0, 1, 0, 0, 0, 0)$
- $\tilde{\mathbf{q}}^{17} = (0, 0, 0, 1, 0, 0, 0, 1)$
- $\tilde{\mathbf{q}}^{18} = (0, 0, 0, 1, 0, 0, 1, 0)$
- $\tilde{\mathbf{q}}^{19} = (0, 0, 0, 1, 0, 0, 1, 1)$
- $\tilde{\mathbf{q}}^{20} = (0, 0, 0, 1, 0, 1, 0, 0)$
- $\tilde{\mathbf{q}}^{21} = (0, 0, 0, 1, 0, 1, 0, 1)$
- $\tilde{\mathbf{q}}^{22} = (0, 0, 0, 1, 0, 1, 1, 0)$
- $\tilde{\mathbf{q}}^{23} = (0, 0, 0, 1, 0, 1, 1, 1)$
- $\tilde{\mathbf{q}}^{24} = (0, 0, 0, 1, 1, 0, 0, 0)$
- $\tilde{\mathbf{q}}^{25} = (0, 0, 0, 1, 1, 0, 0, 1)$
- $\tilde{\mathbf{q}}^{26} = (0, 0, 0, 1, 1, 0, 1, 0)$
- $\tilde{\mathbf{q}}^{27} = (0, 0, 0, 1, 1, 0, 1, 1)$
- $\tilde{\mathbf{q}}^{28} = (0, 0, 0, 1, 1, 1, 0, 0)$
- $\tilde{\mathbf{q}}^{29} = (0, 0, 0, 1, 1, 1, 0, 1)$
- $\tilde{\mathbf{q}}^{30} = (0, 0, 0, 1, 1, 1, 1, 0)$
- $\tilde{\mathbf{q}}^{31} = (0, 0, 0, 1, 1, 1, 1, 1)$
- $\tilde{\mathbf{q}}^{32} = (0, 0, 1, 0, 0, 0, 0, 0)$
- $\tilde{\mathbf{q}}^{33} = (0, 0, 1, 0, 0, 0, 0, 1)$
- $\tilde{\mathbf{q}}^{34} = (0, 0, 1, 0, 0, 0, 1, 0)$
- $\tilde{\mathbf{q}}^{35} = (0, 0, 1, 0, 0, 0, 1, 1)$
- $\tilde{\mathbf{q}}^{36} = (0, 0, 1, 0, 0, 1, 0, 0)$
- $\tilde{\mathbf{q}}^{37} = (0, 0, 1, 0, 0, 1, 0, 1)$
- $\tilde{\mathbf{q}}^{38} = (0, 0, 1, 0, 0, 1, 1, 0)$
- $\tilde{\mathbf{q}}^{39} = (0, 0, 1, 0, 0, 1, 1, 1)$
- $\tilde{\mathbf{q}}^{40} = (0, 0, 1, 0, 1, 0, 0, 0)$
- $\tilde{\mathbf{q}}^{41} = (0, 0, 1, 0, 1, 0, 0, 1)$
- $\tilde{\mathbf{q}}^{42} = (0, 0, 1, 0, 1, 0, 1, 0)$
- $\tilde{\mathbf{q}}^{43} = (0, 0, 1, 0, 1, 0, 1, 1)$
- $\tilde{\mathbf{q}}^{44} = (0, 0, 1, 0, 1, 1, 0, 0)$
- $\tilde{\mathbf{q}}^{45} = (0, 0, 1, 0, 1, 1, 0, 1)$
- $\tilde{\mathbf{q}}^{46} = (0, 0, 1, 0, 1, 1, 1, 0)$
- $\tilde{\mathbf{q}}^{47} = (0, 0, 1, 0, 1, 1, 1, 1)$

- 20

- $\tilde{\mathbf{q}}^{242} = (1, 1, 1, 1, 0, 0, 1, 0)$
- $\tilde{\mathbf{q}}^{243} = (1, 1, 1, 1, 0, 0, 1, 1)$
- $\tilde{\mathbf{q}}^{244} = (1, 1, 1, 1, 0, 1, 0, 0)$
- $\tilde{\mathbf{q}}^{245} = (1, 1, 1, 1, 0, 1, 0, 1)$
- $\tilde{\mathbf{q}}^{246} = (1, 1, 1, 1, 0, 1, 1, 0)$
- $\tilde{\mathbf{q}}^{247} = (1, 1, 1, 1, 0, 1, 1, 1)$
- $\tilde{\mathbf{q}}^{248} = (1, 1, 1, 1, 1, 0, 0, 0)$
- $\tilde{\mathbf{q}}^{249} = (1, 1, 1, 1, 1, 0, 0, 1)$
- $\tilde{\mathbf{q}}^{250} = (1, 1, 1, 1, 1, 0, 1, 0)$
- $\tilde{\mathbf{q}}^{251} = (1, 1, 1, 1, 1, 0, 1, 1)$
- $\tilde{\mathbf{q}}^{252} = (1, 1, 1, 1, 1, 1, 0, 0)$
- $\tilde{\mathbf{q}}^{253} = (1, 1, 1, 1, 1, 1, 0, 1)$
- $\tilde{\mathbf{q}}^{254} = (1, 1, 1, 1, 1, 1, 1, 0)$
- $\tilde{\mathbf{q}}^{255} = (1, 1, 1, 1, 1, 1, 1, 1)$

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.
- 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.