Reactive strategies with longer memory

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

1 Formal Model

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

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

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

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

with $T > R > S > P$ 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 p* is a string $h^p = (a^p_{-1}, \ldots, a^p_{-n}) \in \{C, D\}^n$. An entry a^p_{-k} corresponds to player p's action k rounds ago. Let H^p denote the space of all n-histories of player p. Analogously, let H^q as the set of n-histories h^q of player q. Sets H^p and H^q contain $|H^p| = |H^q| = 2^n$ elements each.

A pair $h = (h^p, h^q)$ is called an *n-history of the game*. We use $H = H^p \times H^q$ to denote the space of all such histories. This set contains $|H| = 2^{2n}$ elements.

Memory-n strategies. A *memory-n* strategy is a vector $\mathbf{m} = (m_h)_{h \in H} \in [0,1]^{2n}$. Each entry m_h corresponds to the player's cooperation probability in the next round, depending on the outcome of the previous n rounds. If the two players use memory-n strategies $\mathbf{m}$ and $\mathbf{m}'$, one can represent the interaction as a Markov chain with a $2^{2n} \times 2^{2n}$ transition matrix M. 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 $\mathbf{S}^k = (S^k_h)_{h \in H}$ denote the vector that returns for each h the one-shot payoff that player p obtained k rounds ago,

$$S^k_h = \begin{cases} b-c & \text{if } a^p_{-k} = C \text{ and } a^q_{-k} = C \\ -c & \text{if } a^p_{-k} = C \text{ and } a^q_{-k} = D \\ b & \text{if } a^p_{-k} = D \text{ and } a^q_{-k} = C \\ 0 & \text{if } a^p_{-k} = D \text{ and } a^q_{-k} = D \end{cases} \tag{3}$$

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

$$s_{\mathbf{m},\mathbf{m}'} = \mathbf{v} \cdot \mathbf{S}^1 = \mathbf{v} \cdot \mathbf{S}^2 = \ldots = \mathbf{v} \cdot \mathbf{S}^n. \tag{4}$$

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. The payoff $s_{\mathbf{m}',\mathbf{m}}$ of player q can be defined analogously.

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

Nash Strategies. A strategy $\mathbf{m}$ for player p, is a *Nash strategy*, if player q never receives a payoff higher than that of the mutual cooperation payoff. Irrespective of q's strategy. Namely if,

$$s_{\mathbf{m}',\mathbf{m}} \leq (b - c) \; \forall \; m'. \tag{5}$$

Nice Strategies. A player's strategy is *nice*, if the player is never the first to defect.

Partner Strategies. For player p, a *partner strategy* is a nice strategy such that,

$$s_{\mathbf{m}',\mathbf{m}} < (b - c) \;\Rightarrow\; s_{\mathbf{m},\mathbf{m}'} < (b - c), \quad \textit{and} \tag{6}$$

$$s_{\mathbf{m}',\mathbf{m}} \geq (b - c) \;\Rightarrow\; s_{\mathbf{m}',\mathbf{m}} = s_{\mathbf{m},\mathbf{m}'} = (b - c). \tag{7}$$

irrespective of the co-player's strategy. In other words, partners strive to achieve the mutual cooperation payoff R 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, making them Nash strategies Hilbe et al. [2015]. All partner strategies are Nash strategies, but not all Nash strategies are partner strategies.

%ToDo Why are partner strategies interesting to study?

Previously the work, of [Akin, 2016] characterized all partner strategies for $n = 1$. For higher memory ($n > 1$) a few works [Hilbe et al., 2017] have managed to characterized partner strategies bit only a subset of them because as memory increases analytical results become more difficult to obtain. However, in this work we characterize all partner reactive strategies for $n = 2, n = 3$. We formally introduce reactive strategies and present the results from section 3 onwards. In the next section, we will discuss a series of results for the general case of memory$-n$.

2 An Extension of Akin's Lemma

The work of [Akin, 2016] focuses on the case of memory-one strategies, thus for $n = 1$. A memory-one strategy of player p is the vector $\mathbf{m} = (m_1, m_2, m_3, , m_4)$, and against a co-player $\mathbf{m}'$ the stationary distribution is of $\mathbf{v} = (v_1, v_2, v_3, v_4)$. Akin's lemma states the following,

Lemma 2.1 (Akin's Lemma)**.** Assume that player p uses the memory-one strategy $\mathbf{m} = (m_1, m_2, m_3, m_4)$, and q uses a strategy that leads to a sequence of distributions $\{\mathbf{v}^{(n)}, n = 1, 2, ...\}$ with $\mathbf{v}^{(k)}$ representing the distribution over the states in the k^{th} round of the game. Let $\mathbf{v}$ be the associated stationary distribution, and let $\tilde{\mathbf{m}} = \mathbf{m} - \mathbf{e}_{12}$ where $\mathbf{e}_{12} = (1, 1, 0, 0)$. Then,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbf{v}^{(k)}\cdot\tilde{\mathbf{m}} = 0, \text{ and therefore } \mathbf{v}\cdot\tilde{\mathbf{m}} = 0. \tag{8}$$

$$\mathbf{v}\cdot\tilde{\mathbf{m}} = (m_{CC}-1)v_{CC}+(m_{CD}-1)v_{CD}+m_{DC}v_{DC}+m_{DD}v_{DD}. \tag{9}$$

The interpretation of this lemma is that the player's probabilities p of switching from cooperation to defection and from defection to cooperation are equal. This is due to the fact that player p can only switch from cooperation to defection if they have previously switched from defection to cooperation.

In the following we generalise Akin's Lemma to $n>1$. Before we do so, we provide some further, definition.

One special case of such a memory-n strategy is the *round-k-repeat strategy*. Player p 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 $a^p_{-k}=C$, then $m_h^{k-\text{Rep}}=1$; otherwise $m_h^{k-\text{Rep}}=0$.

With the same method as in [Akin, 2016], one can show *Akin's Lemma*: For each k with $1\le k\le n$, the invariant distribution $\mathbf{v}$ satisfies the following relationship,

$$\mathbf{v}\cdot(\mathbf{m}-\mathbf{m}^{k-\text{Rep}})=\sum_{h\in H} v_h(m_h-m_h^{k-\text{Rep}})=0. \tag{10}$$

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 $\sum_{h\in H} v_h m_h$.

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

$$\mathbf{m}=\alpha\mathbf{S}^{k_1}+\beta\tilde{\mathbf{S}}^{k_2}+\gamma\mathbf{1}+\mathbf{m}^{k-\text{Rep}}, \tag{11}$$

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

$$0=\mathbf{v}\cdot(\mathbf{m}-\mathbf{m}^{k-\text{Rep}})=\mathbf{v}\cdot(\alpha\mathbf{S}^{k_1}+\beta\tilde{\mathbf{S}}^{k_2}+\gamma\mathbf{1})=\alpha s_{\mathbf{m},\mathbf{m}'}+\beta s_{\mathbf{m}',\mathbf{m}}+\gamma. \tag{12}$$

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 (11) yields the strategy

$$m_h=\begin{cases} 1 & \text{if } a^q_{-k}=C \\ 0 & \text{if } a^q_{-k}=D \end{cases} \tag{13}$$

That is, this strategy implements Tit-for-Tat (for $k=1$) or delayed versions thereof (for $k>1$). By Eq. (12), the enforced payoff relationship is $s_{\mathbf{p}}=s_{\mathbf{q}}$ (in particular, these strategies are *partners*).

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. (11) yields the strategy

$$m_h = \begin{cases} 1 & \text{if } a^q_{-k} = C \\ 1 - c/b & \text{if } a^q_{-k} = D \end{cases} \tag{14}$$

That is, the generated strategy is GTFT (if $k=1$), or delayed versions thereof (for $k>1$). By Eq. (12), the enforced payoff relationship is $s_{\mathbf{m}',\mathbf{m}} = b-c$. In particular, these strategies are not *partner strategies*, but they satisfy the notion of being *Nash strategies*.

The two aforementioned results can be summarized as follows:

- Any Tit-for-Tat strategy for any n, including delayed versions for $k > 1$, is considered a partner strategy.
- Any GTFT strategy for any n, including delayed versions for $k > 1$, is considered a partner strategy.

%ToDo Should these results be propositions?

3 Reactive Partner Strategies

A $n-$ *bit reactive strategy* is denoted by a vector $\mathbf{p} = (p_h)_{h \in H^q} \in [0,1]^{2n}$. 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.

By concentrating on this specific set of strategies, we derive a sequence of intriguing results.

To begin, let's introduce some additional notation. Suppose player p adopts are 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 with,

$$\sum_{h \in H} v_h = 1. \tag{15}$$

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

$$v^q_h = \sum_{h^p \in H^p} v_{(h^p, h^q)} \; \forall \; h^q \in H^q. \tag{16}$$

These entries describe how often we observe player q to choose action(s) h^q, in n consecutive rounds (irrespective of the actions of player p). Based on the above notation, we can define player q's average cooperation rate $\rho_{\mathbf{m}}$. Let, H^q_C be the subset of H^q,

$$H^q_C = \{h^q \in H^q : (h^q_{-2}, h^q_{-1}) = (C, C) \;\vee\; (h^q_{-2}, h^q_{-1}) = (C, D)\}, \quad \textit{then} \tag{17}$$

$$\rho_{\mathbf{m}} := \sum_{h \in H^q_C} v^q_h. \tag{18}$$

Similarly, we can express player p's average cooperation rate $\rho_{\mathbf{p}}$ in terms of v^q_h by noting that

$$\rho_{\mathbf{p}} \quad = \quad \sum_{h \in H^q} v^q_h \, p_h. \tag{19}$$

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

$$\begin{aligned} s_{\mathbf{p},\mathbf{m}} &= b\,\rho_{\mathbf{m}} - c\,\rho_{\mathbf{p}} \\ s_{\mathbf{m},\mathbf{q}} &= b\,\rho_{\mathbf{p}} - c\,\rho_{\mathbf{m}}. \end{aligned} \tag{20}$$

3.1 Sufficiency of Self reactive strategies

To characterize all partner n-bit reactive strategies, one would usually need to check against all pure n-memory one strategies McAvoy and Nowak [2019]. However, we demonstrate that when player p employs an n-bit reactive strategy, it is sufficient to check only against $n-$bit self-reactive 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 n-history. Formally, a self-reactive-n strategy is a vector $\tilde{\mathbf{p}} = (\tilde{p}_h)_{h \in H^q} \in [0,1]^2 n$. 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.

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:

$$v_h = u^q_h \ \forall \ h \in H^q. \tag{21}$$

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:

$$p_{DD} < 1-\frac{c}{b} \quad \textit{and} \quad \frac{p_{CD} + p_{DC}}{2} < 1 - \frac{1}{2} \cdot \frac{c}{b}. \tag{22}$$

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,

$$\begin{aligned}
v^q_{CC} &= \sum_{h^p \in H^p} v_{(h^p, CC)} \\
v^q_{CD} &= \sum_{h^p \in H^p} v_{(h^p, CD)} \\
v^q_{DC} &= \sum_{h^p \in H^p} v_{(h^p, DC)} \\
v^q_{DD} &= \sum_{h^p \in H^p} v_{(h^p, DD)}.
\end{aligned} \tag{23}$$

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

$$\rho_{\mathbf{m}} := v^q_{CC} + v^q_{CD} = v^q_{CC} + v^q_{DC}. \tag{24}$$

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

$$v^q_{CD} = v^q_{DC}. \tag{25}$$

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

$$\begin{aligned}
\rho_{\mathbf{p}} &= v^q_{CC}\, p_{CC} + v^q_{CD}\, p_{CD} + v^q_{DC}\, p_{DC} + v^q_{DD}\, p_{DD} \\
&= v^q_{CC}\, p_{CC} + v^q_{CD}\, (p_{CD} + p_{DC}) + v^q_{DD}\, p_{DD}.
\end{aligned} \tag{26}$$

Here, the second equality is due to Eq. (25).

Finally, we note that we trivially have the following relationship (since all probabilities need to add up to one),

$$1 = v^q_{CC} + v^q_{CD} + v^q_{DC} + v^q_{DD} = v^q_{CC} + 2v^q_{CD} + v^q_{DD} \tag{27}$$

After these preparations, we can prove our theorem based on the same method as in Akin [2016].

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

$$\begin{aligned}
0 \quad &\leq \quad s_{\mathbf{m},\mathbf{p}} - (b-c) \\
&\overset{Eq.\ (20)}{=} \quad b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b-c) \\
&\overset{Eqs.\ (24),(26),(27)}{=} \quad b\left(v^q_{CC}p_{CC} + v^q_{CD}(p_{CD}+p_{DC}) + v^q_{DD}p_{DD}\right) - c\left(v^q_{CC} + v^q_{CD}\right) - (b-c)\left(v^q_{CC} + 2v^q_{CD} + v^q_{DD}\right) \\
&= \quad v^q_{CC}\, b\,(p_{CC}-1) + v^q_{CD}\Big(b(p_{CD}+p_{DC}) + c - 2b\Big) + v^q_{DD}\Big(bp_{DD} - (b-c)\Big).
\end{aligned} \tag{28}$$

By assumption (22),

$$p_{CC} = 1, \qquad b(p_{CD}+p_{DC})+c-2b < 0, \qquad bp_{DD} - (b-c) < 0. \tag{29}$$

Because any $v^q_{XY} \geq 0$, inequality (28) can only hold if $v^q_{CD} = v^q_{DD} = 0$, which implies $v^q_{DC} = 0$ because of Eq. (25). But then it follows that $v^q_{CC} = 1$. By Eqs. (24) and (26) it follows that $\rho_{\mathbf{m}} = \rho_{\mathbf{p}} = 1$, and hence $s_{\mathbf{m},\mathbf{p}} = s_{\mathbf{p},\mathbf{m}} = b-c$. □

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,

$$s_{\mathbf{m},\mathbf{p}} \leq (b-c), \tag{30}$$

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 $i \in 1,\ldots,16$. The strategies are as follow,

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

Proof. Let the following payoffs of a nice reactive-two strategy p against the set of pure self-reactive-two strategies,

$$
\begin{aligned}
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = b \times p_{CC} \quad for \quad i \in \{0,2,4,6,8,10,12,14\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_{CD}+p_{DC}+p_{DD})}{3} - \tfrac{c}{3} \quad for \quad i \in \{1,9\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_{CD}+p_{DC}+p_{DD}+1)}{4} - \tfrac{c}{2} \quad for \quad i \in \{3\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_{CD}+p_{DC})}{2} - \tfrac{c}{2} \quad for \quad i \in \{4,5,12,13\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_{CD}+p_{DC}+1)}{3} - \tfrac{2c}{2} \quad for \quad i \in \{6,7\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = b - c \quad for \quad i \in \{8,9,10,11,12,13,14,15\}
\end{aligned}
\tag{31}
$$

Setting expression of Eq. (31) to smaller than $(b-c)$ we get the three following conditions,

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

$$\frac{p_{CD}+p_{DC}+p_{DD}}{3} < 1 - \frac{2c}{3b} \tag{33}$$

$$\frac{p_{CD}+p_{DC}}{2} < 1 - \frac{c}{2b} \tag{34}$$

$$\tag{35}$$

Note that condition (34) is the sum of conditions (33) and (35). Thus, only conditions (33) and (35) 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:

$$\frac{p_{CCD}+p_{CDC}+p_{DCC}}{3} < 1 - \frac{1}{3}\cdot\frac{c}{b} \qquad \frac{p_{CDD}+p_{DCD}+p_{DDC}}{3} < 1 - \frac{2}{3}\cdot\frac{c}{b} \qquad p_{DDD} < 1 - \frac{c}{b} \tag{36}$$

$$\frac{p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC}}{4} < 1 - \frac{1}{2}\cdot\frac{c}{b} \qquad \frac{p_{CDC}+p_{DCD}}{2} < 1 - \frac{1}{2}\cdot\frac{c}{b} \tag{37}$$

Once again, there are two independent proves of Theorem 3.3, and 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,

$$
\begin{aligned}
v^q_{CCC} &= \sum_{h^p \in H^p} v_{(h^p,CCC)} \\
v^q_{CCD} &= \sum_{h^p \in H^p} v_{(h^p,CCD)} \\
v^q_{CDC} &= \sum_{h^p \in H^p} v_{(h^p,CDC)} \\
v^q_{CDD} &= \sum_{h^p \in H^p} v_{(h^p,CDD)} \\
v^q_{DCC} &= \sum_{h^p \in H^p} v_{(h^p,DCC)} \\
v^q_{DCD} &= \sum_{h^p \in H^p} v_{(h^p,DCD)} \\
v^q_{DDC} &= \sum_{h^p \in H^p} v_{(h^p,DDC)} \\
v^q_{DDD} &= \sum_{h^p \in H^p} v_{(h^p,DDD)}.
\end{aligned}
\tag{38}
$$

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

$$\rho_{\mathbf{m}} := v^q_{CCC} + v^q_{CCD} + v^q_{DCC} + v^q_{DCD} \tag{39}$$

Note that the following equalities hold in the case of $n = 3$,

$$v^q_{CCD} = v^q_{DCC} \tag{40}$$
$$v^q_{DDC} = v^q_{CDD} \tag{41}$$
$$v^q_{CCD} + v^q_{DCD} = v^q_{CDC} + v^q_{DDC} \tag{42}$$
$$\tag{43}$$

The average cooperation rate of p's is given by

$$
\begin{aligned}
\rho_{\mathbf{p}} \quad &= \quad v^q_{CCC}\, p_{CCC} + v^q_{CCD}\, p_{CCD} + v^q_{CDC}\, p_{CDC} + v^q_{CDD}\, p_{CDD} + v^q_{DCD}\, p_{DCD} + v^q_{DDC}\, p_{DDC} + v^q_{DDD}\, p_{DDD} \\
&\overset{Eq.\ (41),\ (42)}{=} \quad v^q_{CCC}\, p_{CCC} + v^q_{CCD}\,(p_{CCD} + p_{DCC}) + v^q_{CDC}\, p_{CDC} + v^q_{CDD}\,(p_{CDD} + p_{DDC}) + v^q_{DCD}\, p_{DCD} + v^q_{DDD}\, p_{DDD} \\
&\overset{Eq.\ (43)}{=} \quad v^q_{CCC}\, p_{CCC} + v^q_{CCD}\,(p_{CCD} + p_{CDC} + p_{DCC}) + v^q_{DCD}\,(p_{CDC} + p_{DDC}) v^q_{DDD}\, p_{DDD} + \\
&\qquad v^q_{DDC}\,(p_{CDD} + p_{DCD} + p_{DDC}) + v^q_{CDD}\,(-p_{CDC} - p_{DDC}) + v^q_{DDD}\, p_{DDD} \\
&\overset{Eq.\ (15)}{=} \quad v^q_{CCC}\, p_{CCC} + v^q_{CCD}\,(p_{CCD} + p_{CDC} + p_{DCC}) + v^q_{DCD}\,(p_{CDC} + p_{DDC}) v^q_{DDD}\, p_{DDD} + \\
&\qquad v^q_{DDC}\,(p_{CDD} + p_{DCD} + p_{DDC}) + v^q_{CDD}\,(p_{CCD} + p_{DCC} + p_{DCD} + p_{DDC}) + (v^q_{DDD} + v^q_{CDD})\, p_{DDD}
\end{aligned}
\tag{44}
$$

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

$$
\begin{aligned}
0 &\leq s_{\mathbf{m},\mathbf{p}} - (b-c) \\
&\overset{Eq.~(20)}{=} b\rho_{\mathbf{p}} - c\rho_{\mathbf{m}} - (b-c) \\
&\overset{Eqs.~(24),(26),(27)}{=} b\left(v^q_{CC}p_{CC}+v^q_{CD}(p_{CD}+p_{DC})+v^q_{DD}p_{DD}\right) - c\left(v^q_{CC}+v^q_{CD}\right) - (b-c)\Big(v^q_{CC}+2v^q_{CD}+v^q_{DD}\Big) \\
&= v^q_{CC}\, b\,(p_{CC}-1) + v^q_{CD}\Big(b(p_{CD}+p_{DC})+c-2b\Big) + v^q_{DD}\Big(bp_{DD}-(b-c)\Big).
\end{aligned}
\tag{45}
$$

□

Proof Two. Consider all the pure self-reactive-three strategies, there are a total of 256 of them. These are given in the appendix. regardless, the payoff expressions for each of these strategies against a nice reactive-three strategies can be calculated explicilty. We will 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,

$$
\begin{aligned}
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_4+p_6+p_7+p_8)-c}{4} \quad for \quad i \in \{1,9,33,41,65,73,97,105,129,137,161,169,193,201,225,233\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_2+p_4+p_5+p_7+p_8)-2c}{5} \quad for \quad i \in \{3,7,35,39,131,135,163,167\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_3+p_6)-c}{2} \quad for \quad i \in \{4-7,12-15,20-23,28-31,68-71,76-79,84-87,92-95,132-135,140-143,148-151,156-159,196- \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_2+p_4+p_5+p_7+p_8+1)}{6} - \tfrac{c}{2} \quad for \quad i \in \{11,,15,,43,,47\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_4+p_6+p_7)-c}{3} \quad for \quad i \in \{16,17,24,25,48,49,56,57,80,81,88,89,112,113,120,121,144,145,152,153,176,177,184,185,208,20 \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_2+p_4+p_5+p_7)-2c}{4} \quad for \quad i \in \{18,19,22,23,50,51,54,55,146,147,150,151,178,179,182,183\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_2+p_4+p_5+p_7+1)-3c}{5} \quad for \quad i \in \{26,27,30,31,58,59,62,63\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_2+p_3+p_4+p_5+p_6+p_7+p_8)-3c}{7} \quad for \quad i \in \{37,67,165,195\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_2+p_3+p_4+p_5+p_6+p_7+p_8+1)}{8} - \tfrac{c}{2} \quad for \quad i \in \{45,75\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_2+p_3+p_4+p_5+p_6+p_7)}{6} - \tfrac{c}{2} \quad for \quad i \in \{52,53,82,83,180,181,210,211\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_2+p_3+p_4+p_5+p_6+p_7+1)-4c}{7} \quad for \quad i \in \{60,61,90,91\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_2+p_3+p_5)-2c}{3} \quad for \quad i \in \{96-103,112-119,224-231,240-247\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = \tfrac{b(p_2+p_3+p_5+1)-3c}{4} \quad for \quad i \in \{104-111,120-127\} \\
&s_{\tilde{\mathbf{q}}^i,\mathbf{p}} = (b-c) \quad for \quad i \in [128,255]
\end{aligned}
\tag{46}
$$

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

$$p_8 \leq 1-\frac{c}{b} \tag{47}$$

$$p_4+p_6+p_7+p_8 \leq 4-\frac{3c}{b} \tag{48}$$

$$p_2+p_4+p_5+p_7+p_8 \leq 5-\frac{3c}{b} \tag{49}$$

$$p_3+p_6 \leq 2-\frac{c}{b} \tag{50}$$

$$p_4+p_6+p_7 \leq 3-\frac{2c}{b} \tag{51}$$

$$p_2+p_4+p_5+p_7 \leq 4-\frac{2c}{b} \tag{52}$$

$$p_2+p_3+p_4+p_5+p_6+p_7+p_8 \leq 7-\frac{4c}{b} \tag{53}$$

$$p_2+p_3+p_4+p_5+p_6+p_7 \leq 6-\frac{3c}{b} \tag{54}$$

$$p_2+p_3+p_5 \leq 3-\frac{c}{b} \tag{55}$$

Note that only conditions are unique. The following can be derived from the sums of two or more of these conditions.

□

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 $\mathbf{r} = (r_i)_{i\in[0,dots,n]}$, where the entries r_i indicate 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 (22). 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,

$$r_1 < 1 - \frac{1}{2}\cdot\frac{c}{b} \ \ \textit{and} \ \ r_0 < 1-\frac{c}{b}. \tag{56}$$

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 (36). 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,

$$r_2 < 1 - \frac{1}{3}\cdot\frac{c}{b}, \quad r_1 < 1 - \frac{2}{3}\cdot\frac{c}{b} \ \ \textit{and} \ \ r_0 < 1-\frac{c}{b}. \tag{57}$$

In the case of counting reactive strategies, we observe a pattern in the conditions they must satisfy to be partner strategies. We show that for an $n-$bit counting reactive strategy to be a partner strategy, the strategy's entries must satisfy the conditions:

$$
\begin{aligned}
r_n =& 1 \\
r_{n-1} \leq& 1 - \frac{(n-1)}{n} \times \frac{c}{b} \\
r_{n-2} \leq& 1 - \frac{(n-2)}{n} \times \frac{c}{b} \\
\vdots& \\
r_0 \leq& 1 - \frac{c}{b}
\end{aligned}
$$

$$
\begin{aligned}
H_k^q &= \{h^q \in H^q : \ |A(h^q)| = k\}, \quad for \\
A(h^q) &= \{a^q \in h^q : \ a^q = C\}
\end{aligned}
$$

$$
\rho_{\mathbf{p}} = v^q_{C\dots C}\, r_n \ + \sum_{k=1}^{n-1} r_{n-k} \sum_{h \in H_k^q} v_h^q + \ v^q_{D\dots D} r_0 \tag{58}
$$

4 Figures

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.

C. Hilbe, L. A. Martinez-Vaquero, K. Chatterjee, and M. A. Nowak. Memory-n strategies of direct reciprocity. *Proceedings of the National Academy of Sciences*, 114(18):4715–4720, 2017.

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.

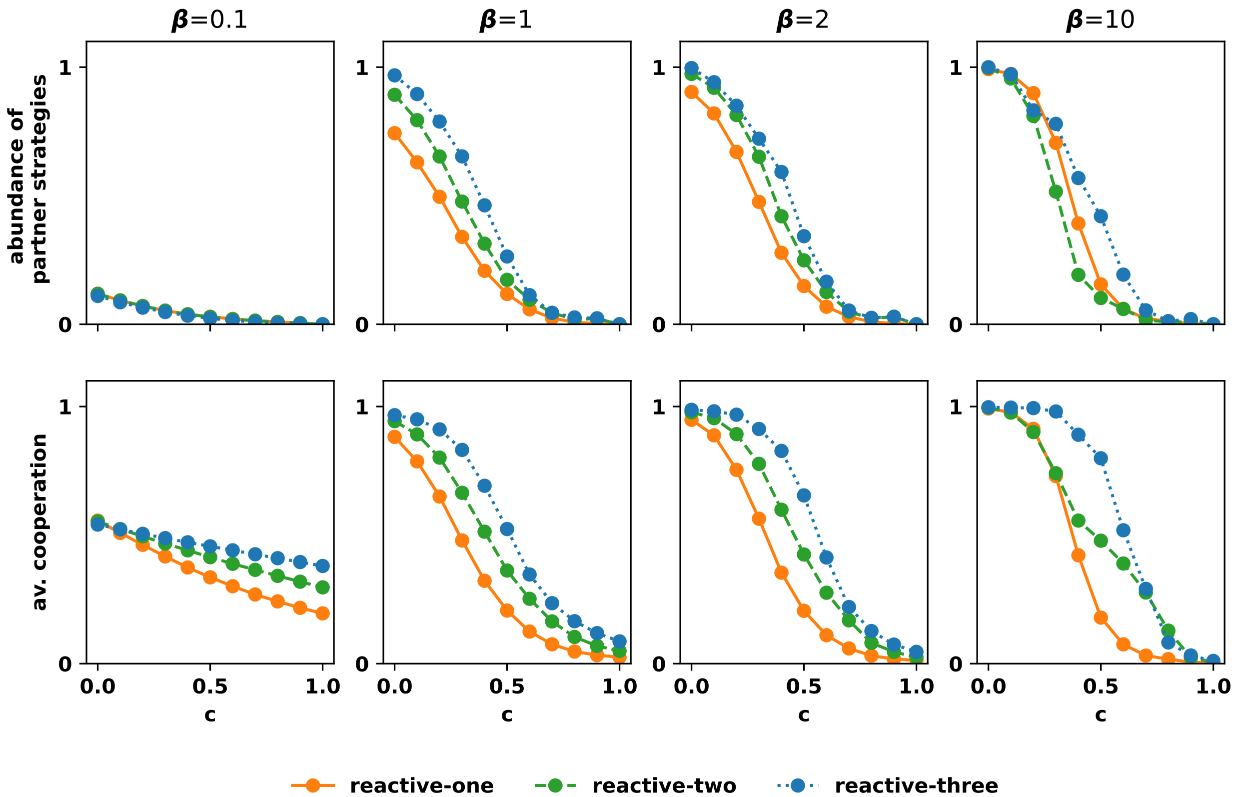

Figure 1: The abundance of partner strategies for $n = 1, 2, 3$ and $b = 1, c = 0.5$.

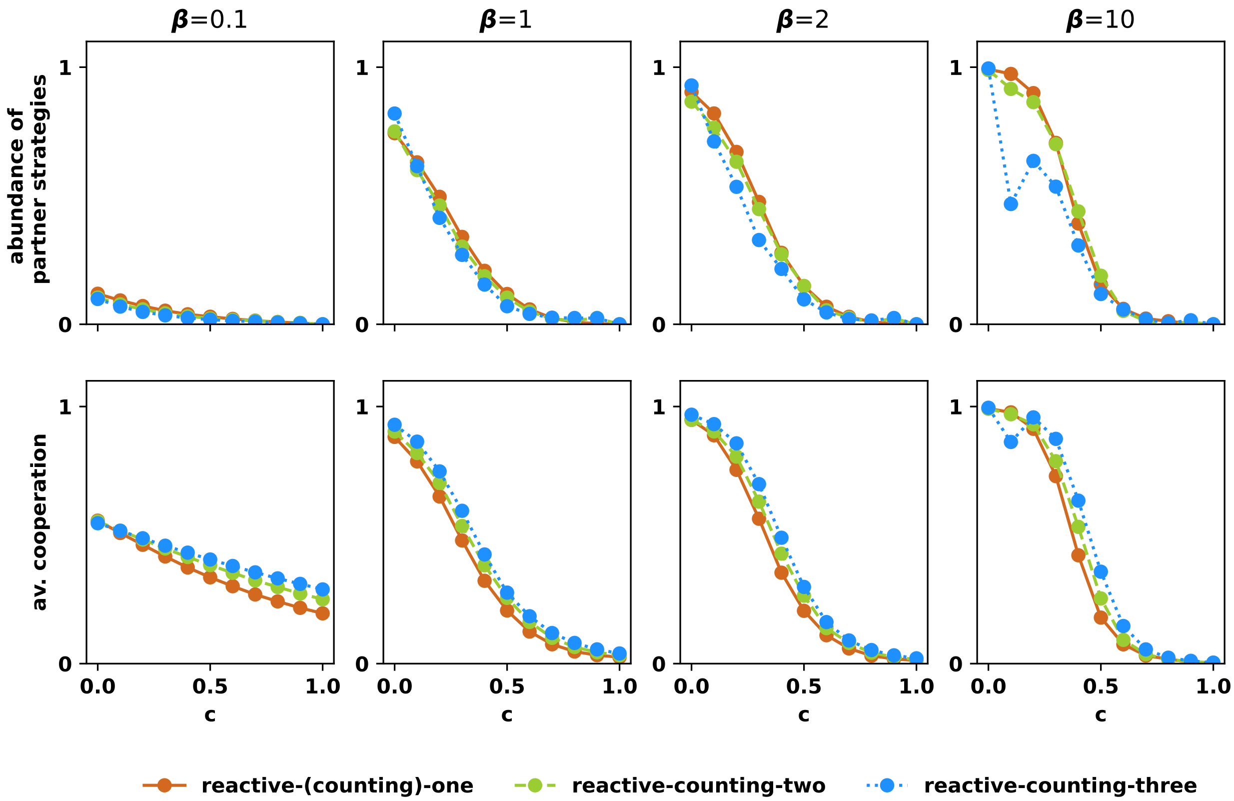

Figure 2: The abundance of partner counting strategies for $n = 1, 2, 3$ and $b = 1, c = 0.5$.

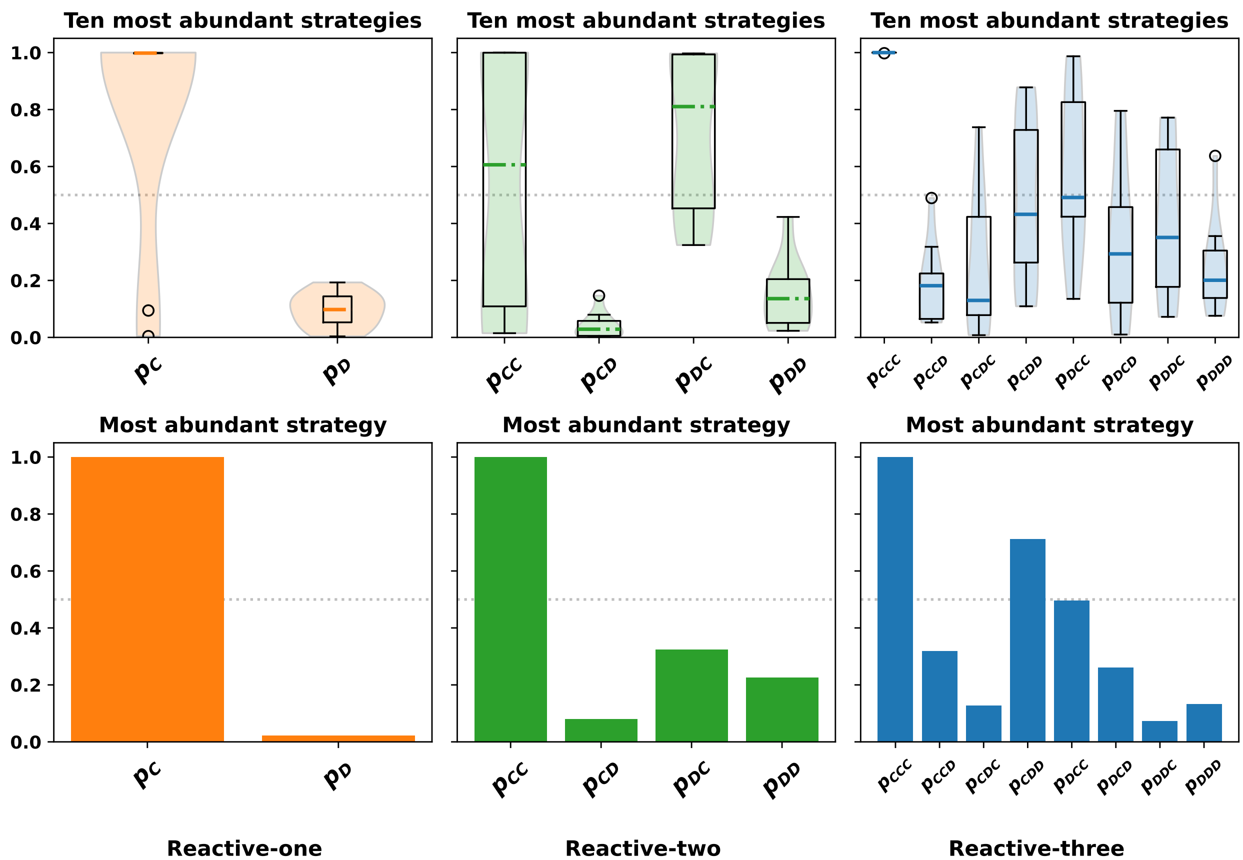

Figure 3: The most abundant reactive-n strategies for $n = 1, 2, 3$ and $b = 1, c = 0.5, \beta = 1$.

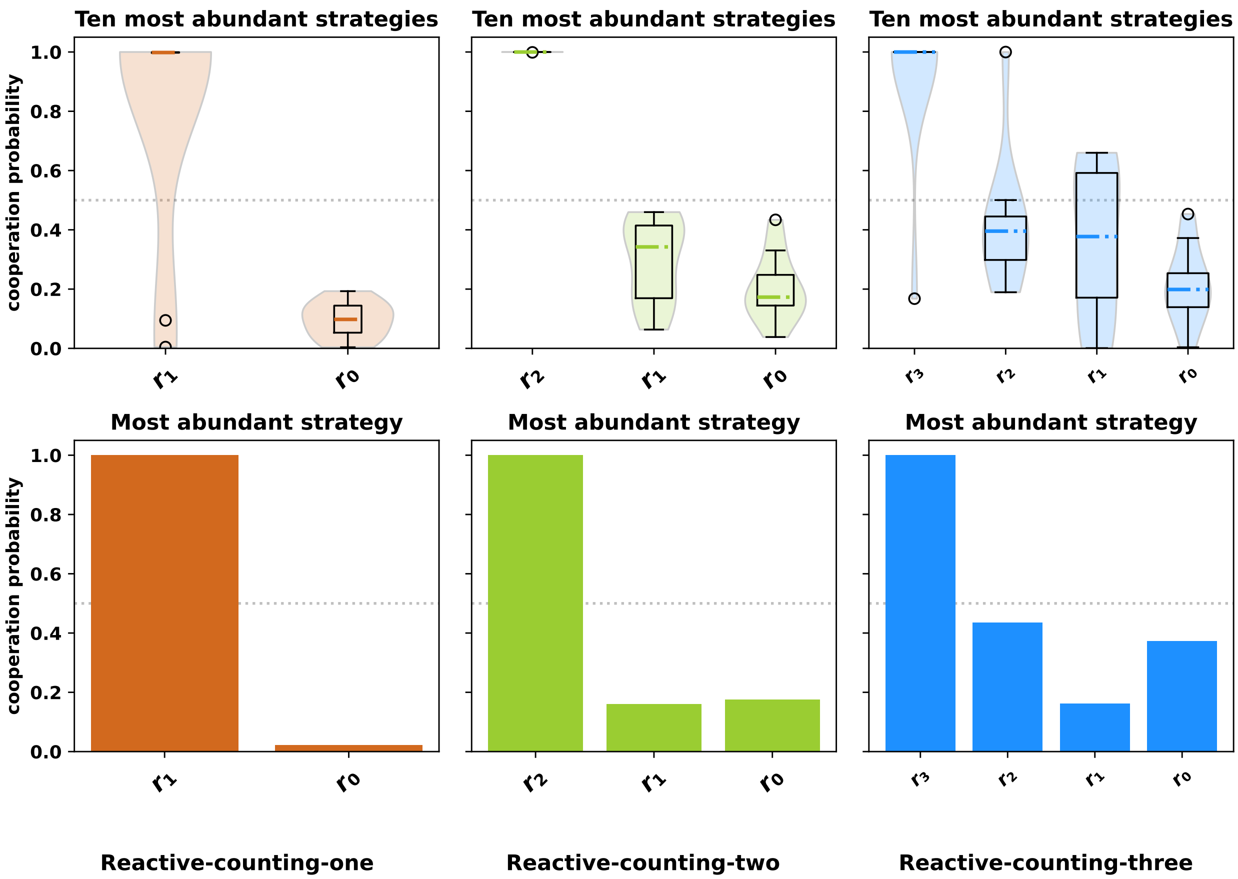

Figure 4: The most abundant reactive-counting-n strategies for $n = 1, 2, 3$ and $b = 1, c = 0.5, \beta = 1$.