SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

August 2026, Vol. 69, Iss. 8, 182203:1–182203:15
https://doi.org/10.1007/s11432-025-4945-7

Evolutionary game dynamics under adaptive coordination

Feipeng ZHANG[1,2], Te WU[1*] & Long WANG[3*]

[1]*Center for Complex Systems, Xidian University, Xi'an 710071, China*
[2]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong 999077, China*
[3]*Center for Systems and Control, College of Engineering, Peking University, Beijing 100871, China*

Received 19 September 2025/Revised 11 November 2025/Accepted 17 December 2025/Published online 16 July 2026

Abstract Decision-making in natural and artificial systems is often shaped by past interactions, as individuals adjust their behavior accordingly. A persistent challenge lies in identifying successful strategies and understanding how memory influences their performance, especially when individuals' optimal actions conflict with collective outcomes. Most theoretical studies in evolutionary dynamics have focused on memory-1 strategies, and recent studies indicate that extending memory can improve strategies' performance. Among these, all-or-none (AoN_K) strategies perform well at short memory. Our analysis reveals, however, that their effectiveness deteriorates as memory extends, showing the inherent limitations of sole coordination. Building on this, we propose the adaptive coordination strategy, combining coordination and tolerance to enhance adaptability and stability. Our theoretical analysis precisely characterizes its behavior and key properties. This strategy outperforms classic memory-1 strategies and achieves optimal outcomes, providing a framework for understanding complex history-based behaviors beyond traditional memory-n models.

Keywords evolutionary game dynamics, adaptive coordination, decision-making, strategies, adaptability

Citation Zhang F P, Wu T, Wang L. Evolutionary game dynamics under adaptive coordination. Sci China Inf Sci, 2026, 69(8): 182203, https://doi.org/10.1007/s11432-025-4945-7

1 Introduction

Individuals often adopt conditional strategies, adjusting their future actions based on past interactions to better adapt to their environment [1–6]. The study of such history-dependent decision-making spans multiple domains, from microbial dynamics and multi-agent learning algorithms to the widespread presence of altruistic behaviors in human societies [7–14]. The definition of a winning strategy depends on the criterion individuals aim to optimize, with personal success being the most natural and immediate objective [15–17]. However, when two self-interested individuals interact, their pursuit of personal gain can result in outcomes that are detrimental to both, especially in the presence of conflicts of interest. This tension is well captured by the repeated prisoner's dilemma [18, 19], a central model in evolutionary game theory. In each round, two players independently choose whether to cooperate or defect. While unilateral defection offers a higher immediate reward, cooperation yields greater collective benefit. A persistent challenge, therefore, lies in identifying strategies that can reconcile individual incentives with collective welfare in such settings [20–22].

In evolutionary dynamics, the mechanism by which conditional strategies sustain cooperation within populations is known as direct reciprocity [17, 23–25]. With few exceptions [16, 26, 27], most evolutionary models have focused on simplified subjects who are restricted to a narrow set of strategies or who only remember the outcome of the most recent round. Within this framework, several memory-1 strategies have been identified as successful, including tit-for-tat (TFT) [28], win-stay lose-shift ($WSLS$) [14], and the class of generous zero-determinant (ZD) strategies [29, 30]. Once memory extends beyond a single round, the size of the strategy space increases explosively. For instance, there are 65536 pure memory-2 strategies and around 1.84×10^{19} pure memory-3 strategies. Among these, the all-or-none (AoN_K) strategies perform particularly well in evolutionary simulations [31, 32], highlighting that longer memory can, in some cases, suffice for the emergence of cooperation. Nevertheless, it remains unclear whether such strategies retain their effectiveness at extended memory lengths. Moreover, the exponential growth of the strategy space quickly renders exhaustive evolutionary exploration computationally infeasible, leaving the potential of longer-memory strategies only partially understood.

* Corresponding author (email: wute@pku.edu.cn, longwang@pku.edu.cn)

A natural question is how individuals make decisions when they can rely on more extensive history than just the most recent outcomes [33–35]. Our theoretical analysis shows that once memory extends beyond a certain length, AoN_K can no longer sustain maximal collective payoff, reflected in a decline of cooperation rates. This finding demonstrates that AoN_K does not provide a satisfactory solution under extended memory. The limitations of AoN_K suggest that coordination alone is insufficient to maintain cooperation across longer time horizons. While the capacity to align actions remains essential, excessive rigidity makes strategies vulnerable to noise. These insights call for more flexible mechanisms that combine coordination with tolerance, allowing individuals to adapt while preserving cooperative stability. To address this challenge, we introduce the adaptive coordination (*AdaCo*) strategy, which integrates both principles to achieve robustness and sustain high levels of cooperation under extended memory.

Within the framework of evolutionary dynamics [36–39], we combine theoretical analysis with numerical simulations to evaluate the performance of *AdaCo* across a wide range of evolutionary settings. In these environments, *AdaCo* competes directly with classic strategies and consistently demonstrates superior robustness. Compared to AoN_K, *AdaCo* is more resilient to implementation errors, a feature that is crucial for sustaining cooperation. It also resists invasion by unconditional defectors: defectors cannot exploit *AdaCo*, while reciprocity among *AdaCo* players reinforces cooperation, leading to their eventual dominance. As a result, populations evolve toward high levels of cooperation. Moreover, *AdaCo* remains competitive when pitted against individual classic strategies or when multiple strategies coexist, further underscoring its adaptability. Together, these findings show that *AdaCo* outperforms traditional memory-1 strategies and extends the scope of cooperation beyond what can be explained by existing memory-n frameworks, offering new perspectives on the evolutionary stability of history-based behaviors.

2 Preliminaries and problem formulation

We consider group decision-making in the framework of repeated cooperative dilemmas, where N individuals independently choose to cooperate (C) or defect (D) in each round. For two players, the prisoner's dilemma serves as the canonical model for studying cooperation. Mutual cooperation (CC) yields a reward R, while mutual defection (DD) results in a punishment P. If players choose different actions (CD or DC), the defector receives the temptation payoff T and the cooperator receives the sucker's payoff S. The payoff parameters satisfy $T > R > P > S$ and $T + S < 2R$, ensuring that defection is the unique Nash equilibrium and that mutual cooperation is preferable to alternating cooperation and defection over consecutive rounds. For multi-player dilemmas, the public goods game, which is also referred to as the multi-player prisoner's dilemma, captures the essential tension between individual and collective interests. Other well-studied extensions include the volunteer dilemma [40], the threshold public goods game [41], and the public goods game with risk [41].

Repeated games allow players to adjust their actions based on the history of interactions. In this study, we focus on infinitely repeated games in noisy environments. Specifically, when a player intends to cooperate, the action is implemented correctly with probability $1-\varepsilon$, while with probability ε an error occurs and defection is implemented instead; the same applies when a player intends to defect. In such settings, players accumulate information about past outcomes over infinitely many rounds of the prisoner's dilemma, which enables history-dependent strategies. In infinitely repeated games, strategies can in principle be arbitrarily complex, as players may condition their actions on the entire history of past interactions.

To render the problem analytically and computationally tractable, a reasonable assumption is that players possess finite memory. A memory-n strategy specifies a player's action based on the outcomes of the most recent n rounds. In the repeated prisoner's dilemma, there are 2^{2n} possible sequences of outcomes, and a memory-n strategy p can be represented as

$$p = (p_{\substack{CC\ldots CC\\CC\ldots CC}}, p_{\substack{CC\ldots CC\\CC\ldots CD}}, \ldots, p_{\substack{DD\ldots DD\\DD\ldots DD}}), \tag{1}$$

where each entry gives the probability of cooperating conditional on the corresponding outcome. The first row of the subscript records the focal player's past actions, and the second row records those of the opponent, ordered from the most distant to the most recent. If all entries are either 0 or 1, the strategy is pure; otherwise, it is stochastic. The initial n actions, before a full history is available, have negligible effects on long-run payoffs under infinite repetition with noise.

Among memory-1 or memory-2 strategies, the class of AoN_K stands out, as they combine mutual cooperation, error correction, and retaliation against defectors. Evolutionary simulations have shown that AoN_K can dominate among short-memory strategies. Yet, it remains unclear whether their advantage persists when longer memory lengths are considered, a question that motivates our present analysis.

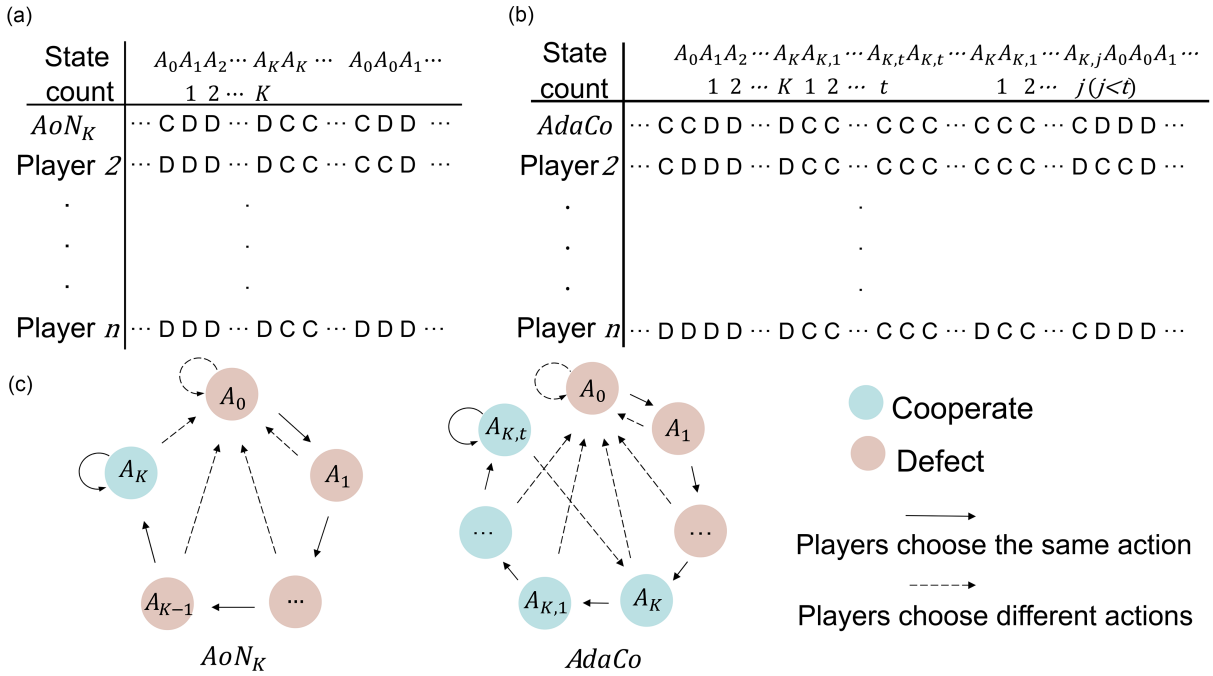

Figure 1 (Color online) Illustration of AoN_K and $AdaCo$. (a) Past I rounds ($I = 1, 2, 3, \dots$) are coordinated if all players choose the same action in each round. AoN_K cooperates if the past K rounds are coordinated; otherwise, it defects. K is the cooperation threshold and the minimum memory length required. The past K-round outcomes define $K + 1$ states A_i ($i = 0, 1, \dots, K$), where A_i indicates i consecutive coordinated rounds, and A_0 represents a recent uncoordinated round. After one uncoordinated round, AoN_K retaliates by defecting for the next K rounds. (b) $AdaCo$ builds on AoN_K by combining the coordination principle with a tolerance mechanism. An $AdaCo$ player with cooperation threshold K and tolerance t behaves like AoN_K until K consecutive coordinated rounds occur. Upon reaching K coordinated rounds, the player cooperates and counts consecutive A_K states. Once this count reaches t, the player tolerates a single uncoordinated round, returning to state A_K rather than A_0, and resumes counting. If an uncoordinated round occurs before reaching t, the state resets to A_0. (c) State-to-state transitions and decision-making for AoN_K and $AdaCo$. State $A_{k,j}$ ($j \in 1, 2, \cdots, t$) indicates that the state A_K has lasted for j rounds. Following a coordinated round, the state A_K transitions to $A_{K,1}$, states $A_{K,j}$ transition to $A_{K,j+1}$ (where $j < t$), and state $A_{K,t}$ remains in its current state. Observing additional t rounds, say observation phase, improves $AdaCo$'s ability to distinguish opponents, especially between defectors and co-species. The parameter t represents the tolerance level of $AdaCo$. The smaller t, the more forgivable. As t tends to infinity, $AdaCo$ with cooperation threshold K is equivalent to AoN_K.

3 Main results

3.1 AoN_K and $AdaCo$ with arbitrary cooperation threshold K

As illustrated in Figure 1, AoN_K prescribes cooperation only if the outcomes of the preceding K rounds are perfectly coordinated, where K denotes the cooperation threshold. Thus, K also represents the minimum memory length required for implementing AoN_K. To systematically examine their performance, we consider AoN_K with arbitrary thresholds K. Although interactions between two AoN_K players can, in principle, be represented by a finite-state Markov chain, the conventional analytical methods developed for general memory-n strategies become increasingly impractical as the memory length n grows, due to the rapidly expanding number of possible states.

To address this challenge, we develop a unified analytical framework for the theoretical analysis of both AoN_K and $AdaCo$, applicable across a wide range of threshold values and interaction settings. This framework substantially reduces the size of the state space, making it feasible to analyze strategies with longer memory. As a result, the interaction dynamics of AoN_K and $AdaCo$ can be derived in a clear and tractable manner.

Proposition 1. Consider an infinitely repeated game among N players, where in each round every player chooses either to cooperate or to defect. Let $\varepsilon > 0$ represent the probability of a player making an error in any given round. Suppose all N players use AoN_K. Then, the group's average cooperation rate is

$$\rho_{AoN_K} = \varepsilon + (1 - 2\varepsilon)\left[(1-\varepsilon)^N + \varepsilon^N\right]^K. \tag{2}$$

In the case of the prisoner's dilemma, the payoff for each player can be expressed as

$$\pi(AoN_K, AoN_K) = \left\{\varepsilon^2 + (1-2\varepsilon)\left[\varepsilon^2 + (1-\varepsilon)^2\right]^K\right\}R + \left\{(1-\varepsilon)^2 - (1-2\varepsilon)\left[\varepsilon^2 + (1-\varepsilon)^2\right]^K\right\}P + \varepsilon(1-\varepsilon)(T+S). \tag{3}$$

Proof. We categorize the outcomes of the past K steps into $K+1$ states $A_0, A_1, \ldots, A_K$ based on the number of latest consistent rounds. A_i means all players chose the same action in each of the most recent i rounds, while A_0 means the most recent round was inconsistent. Let x_i be the probability of being in state A_i and $q = \varepsilon^N + (1-\varepsilon)^N$. Then we have $x_i = qx_{i-1}$ for $i = 1, \ldots, K-1$, and $x_K = qx_{K-1} + qx_K$. Given the normalization $\sum_{i=0}^{K} x_i = 1$, we obtain $x_i = (1-q)q^{i-1}$ for $i = 1, \ldots, K-1$, and $x_K = q^K$.

Now, let us calculate the average cooperation rate of the interaction between these N AoN_K players. By definition, in states 0 to $K-1$, the player's cooperation rate is $C_N^N\varepsilon^N + \frac{N+1}{N}C_N^{N-1}\varepsilon^{N-1}(1-\varepsilon) + \cdots + 0 \cdot C_N^0(1-\varepsilon)^N$ and in state K, the player's cooperation rate is $C_N^N(1-\varepsilon)^N + \frac{N-1}{N}C_N^{N-1}(1-\varepsilon)^{N-1}\varepsilon + \cdots + 0 \cdot C_N^0\varepsilon^N$. So

$$\begin{aligned}\rho_{AoN_K} = &\sum_{i=0}^{K-1} x_i \left[C_N^N\varepsilon^N + \frac{N-1}{N}C_N^{N-1}\varepsilon^{N-1}(1-\varepsilon) + \cdots + 0 \cdot C_N^0(1-\varepsilon)^N\right] \\ &+ x_K \left[C_N^N(1-\varepsilon)^N + \frac{N-1}{N}C_N^{N-1}(1-\varepsilon)^{N-1}\varepsilon + \cdots + 0 \cdot C_N^0\varepsilon^N\right].\end{aligned}$$

Similarly, we can derive the expected payoff when it comes to the two-player prisoner's dilemma.

As $\varepsilon \to 0$, ρ_{AoN_K} approaches one, indicating near-perfect cooperation. However, cooperation rapidly deteriorates in large groups or under high thresholds (long memories). This is because AoN_K retaliates for at least K rounds after a single defection, a mechanism that deters exploitation but simultaneously renders mutual cooperation highly vulnerable to errors. Hence, while AoN_K succeeds under short memory, its error-sensitivity undermines its ability to sustain cooperation in extended-memory settings, highlighting the need for more robust designs.

The error-sensitivity of AoN_K highlights the need for more resilient mechanisms. Motivated by the strong resistance of high-threshold strategies against defectors, we introduce the *AdaCo* strategy, which integrates the coordination principle of AoN_K with an additional tolerance component. As shown in Figure 1, an *AdaCo* player keeps track of the number of consecutive coordinated rounds. When in state A_i ($i \in \{0, 1, \ldots, K\}$), the player behaves like AoN_K: upon reaching K coordinated rounds from A_0, cooperation begins. Unlike AoN_K, however, once t consecutive coordinated rounds are achieved, *AdaCo* tolerates a single uncoordinated round by returning to state A_K rather than resetting to A_0. In this sense, t specifies the tolerance threshold: smaller values of t correspond to greater tolerance.

Proposition 2. Consider the same game parameters as in Proposition 1. Suppose all N players use *AdaCo* with cooperation threshold K and tolerance t. Then, the average cooperation rate of the group is

$$\rho_{AdaCo} = x_0 \left[\frac{1-q^K}{1-q}\varepsilon + \frac{q^K}{(1-q^t)(1-q)}(1-\varepsilon)\right], \tag{4}$$

where $q = \varepsilon^N + (1-\varepsilon)^N$ and $x_0 = \frac{(1-q)(1-q^t)}{(1-q^K)(1-q^K)+q^K}$. In the two-person prisoner's dilemma, each player's payoff is

$$\begin{aligned}\pi(AdaCo, AdaCo) = x_0 \Bigg\{&\left[\frac{1-q^K}{1-q}\varepsilon^2 + \frac{q^K}{(1-q^t)(1-q)}(1-\varepsilon)^2\right] R \\ &+ \left[\frac{1-q^K}{1-q}(1-\varepsilon)^2 + \frac{q^K}{(1-q^t)(1-q)}\varepsilon^2\right] P + \varepsilon(1-\varepsilon)(T+S)\Bigg\}.\end{aligned} \tag{5}$$

Proof. Based on the states given in Proposition 1, we add t states $(A_{K,1}, A_{K,2}, \cdots, A_{K,t})$. Let x_{K+i} ($i = 1, 2, \cdots, t$) denote the probability that a player is in state $A_{K,i}$. According to *AdaCo* rules, the probability x_i satisfies the following equations: $x_i = q \cdot x_{i-1}$ for $i = 0, 1, \cdots, K$; $x_i = q \cdot x_{i-1}$ for $i = K+1, K+2, \cdots, K+t-1$; $x_K = q \cdot x_{K-1} + (1-q)x_{K+t}$; and $x_{K+t} = qx_{K+t-1} + qx_{K+t}$. Given that the probabilities sum to one, we have $1 = \sum_{i=0}^{K+t} x_i = x_0\frac{1-q^K}{1-q} + \frac{x_K}{1-q}$. By solving the above equations, we get

$$x_i = \begin{cases} \frac{q^i(1-q)(1-q^t)}{(1-q^K)(1+q^t)+q^K}, & i = 0, 1, \cdots, K-1, \\ \frac{q^i(1-q)(1-q^t)}{(1-q^K)(1-q^t)^2+(1-q^t)q^K}, & i = K, K+1, \cdots, K+t-1, \\ \frac{q^{K+t}(1-q^t)}{(1-q^K)(1-q^t)^2+(1-q^t)q^K}, & i = K+t. \end{cases}$$

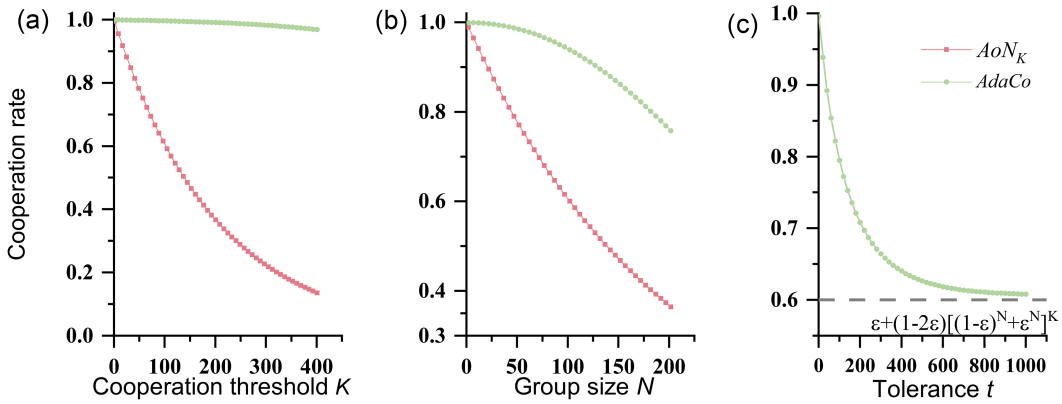

Figure 2 (Color online) Cooperation rate for AoN_K and $AdaCo$ for interactions between co-species. For a strategy to qualify as cooperative, it must not only support mutual cooperation when adopted by all players, but also prevent accidental errors from cascading into persistent defection. Propositions 1 and 2 allow us to compute the resulting cooperation rates for the two strategies across parameter settings. (a) For the same group size, the cooperation rate drastically declines as the cooperation threshold K increases. However, $AdaCo$ groups sustain a high level of cooperation for a wide range of K. (b) The cooperation dilemma becomes more challenging as the group size increases. Even so, $AdaCo$ players still maintain a cooperation level at >75% for group sizes as large as 200. However, the cooperation rate dramatically drops as the group size expands. (c) Cooperation rate for different levels of tolerance, where $K = 100$. $AdaCo$ has limitations in maintaining cooperation when the group size or memory surpasses a certain threshold. A high cooperation rate is difficult to maintain for $AdaCo$ of a low level of forgiveness. Parameters in (a) group size $N = 5$ and $t = 1$, (b) $K = 5$ and $t = 1$, and (c) $K = 100$ and $N = 5$.

Accordingly, the group's average cooperation rate is

$$\rho_{AdaCo} = \sum_{i=0}^{K-1} x_i \left[C_N^N \varepsilon^N + \frac{N-1}{N} C_N^{N-1} \varepsilon^{N-1}(1-\varepsilon) + \cdots + 0 \cdot C_N^0 (1-\varepsilon)^N \right] + \sum_{i=K}^{K+t} x_i \left[C_N^N (1-\varepsilon)^N + \frac{N-1}{N} C_N^{N-1} (1-\varepsilon)^{N-1} \varepsilon + \cdots + 0 \cdot C_N^0 \varepsilon^N \right].$$

Similarly, the expected payoff for the two-player prisoner's dilemma can be derived.

Compared to homogeneous populations of AoN_K players, cooperation levels among $AdaCo$ players are substantially higher across a wide range of K and N, confirming the role of tolerance in mitigating error sensitivity and stabilizing cooperation. Increasing tolerance (smaller t) further promotes cooperation, while in the limit of low tolerance (large t), $AdaCo$ reduces to the behavior of AoN_K (see Figure 2).

3.2 Pairwise interaction analysis

While examining the self-play of a strategy offers valuable insight into its ability to sustain cooperation, it provides only a partial assessment. To obtain a more comprehensive evaluation, we further analyze the pairwise interactions of AoN_K and $AdaCo$ with other representative strategies in the repeated prisoner's dilemma. Following Axelrod's tournament setting, we adopt the canonical payoff parameters $(T, R, P, S) = (5, 3, 1, 0)$ and fix the implementation error at $\varepsilon = 0.1\%$.

Regarding competition between various AoN_K, we can also employ comparable methods to obtain the corresponding payoffs and cooperation rates.

Proposition 3. Consider the same game parameters as in Proposition 1. Suppose that two players use different AoN_K, denoted by AoN_{K_a} and AoN_{K_b} respectively, and that $K_a > K_b$. Then the average cooperation rate is

$$\rho(AoN_{K_a}, AoN_{K_b}) = \varepsilon x_0 \frac{1-q^{K_b}}{1-q} + \frac{1}{2} x_0 q^{K_b} [1-(1-q)^{K_a-K_b}] + (1-\varepsilon) x_0 q^{K_b} (1-q)^{K_a-K_b-1}, \tag{6}$$

where $q = \varepsilon^2 + (1-\varepsilon)^2$ and $x_0 = \left[\frac{1-q^{K_b}}{1-q} + q^{K_b-1} - q^{K_b-1}(1-q)^{K_a-K_b} + q^{K_b}(1-q)^{K_a-K_b-1} \right]^{-1}$.

And the corresponding payoffs are

$$\begin{aligned}\pi_{AoN_{K_a}} = &\left[\frac{1-q^{K_b}}{1-q}\varepsilon^2 + (q^{K_b-1} - q^{K_b-1}(1-q)^{K_a-K_b})\varepsilon(1-\varepsilon) + q^{K_b}(1-q)^{K_a-K_b-1}(1-\varepsilon)^2\right]x_0R\\ &+\left[\frac{1-q^{K_b}}{1-q}\varepsilon(1-\varepsilon) + (q^{K_b-1} - q^{K_b-1}(1-q)^{K_a-K_b})(1-\varepsilon)^2 + q^{K_b}(1-q)^{K_a-K_b-1}\varepsilon(1-\varepsilon)\right]x_0T\\ &+\left[\frac{1-q^{K_b}}{1-q}\varepsilon(1-\varepsilon) + (q^{K_b-1} - q^{K_b-1}(1-q)^{K_a-K_b})\varepsilon^2 + q^{K_b}(1-q)^{K_a-K_b-1}\varepsilon(1-\varepsilon)\right]x_0S\\ &+\left[\frac{1-q^{K_b}}{1-q}(1-\varepsilon)^2 + (q^{K_b-1} - q^{K_b-1}(1-q)^{K_a-K_b})\varepsilon(1-\varepsilon) + q^{K_b}(1-q)^{K_a-K_b-1}\varepsilon^2\right]x_0P\end{aligned} \tag{7}$$

and

$$\begin{aligned}\pi_{AoN_{K_b}} = &\left[\frac{1-q^{K_b}}{1-q}\varepsilon^2 + (q^{K_b-1} - q^{K_b-1}(1-q)^{K_a-K_b})\varepsilon(1-\varepsilon) + q^{K_b}(1-q)^{K_a-K_b-1}(1-\varepsilon)^2\right]x_0R\\ &+\left[\frac{1-q^{K_b}}{1-q}\varepsilon(1-\varepsilon) + (q^{K_b-1} - q^{K_b-1}(1-q)^{K_a-K_b})\varepsilon^2 + q^{K_b}(1-q)^{K_a-K_b-1}\varepsilon(1-\varepsilon)\right]x_0T\\ &+\left[\frac{1-q^{K_b}}{1-q}\varepsilon(1-\varepsilon) + (q^{K_b-1} - q^{K_b-1}(1-q)^{K_a-K_b})(1-\varepsilon)^2 + q^{K_b}(1-q)^{K_a-K_b-1}\varepsilon(1-\varepsilon)\right]x_0S\\ &+\left[\frac{1-q^{K_b}}{1-q}(1-\varepsilon)^2 + (q^{K_b-1} - q^{K_b-1}(1-q)^{K_a-K_b})\varepsilon(1-\varepsilon) + q^{K_b}(1-q)^{K_a-K_b-1}\varepsilon^2\right]x_0P,\end{aligned} \tag{8}$$

respectively.

Proof. Since $K_a > K_b$, the minimum number of states that completely describe the process of the game between the strategies is $K_a + 1$. We first need to calculate the value of x_i. According to the definition of AoN_K, we obtain $x_i = qx_{i-1}$ for $i = 1, \ldots, K_b$, $x_i = (1-q)x_{i-1}$ for $i = K_b + 1, \ldots, K_a - 1$, and $x_{K_a} = (1-q)x_{K_a-1} + qx_{K_a}$. Because $\sum x_i = 1$, we get $x_0 = \left[\frac{1-q^{K_b}}{1-q} + q^{K_b-1} - q^{K_b-1}(1-q)^{K_a-K_b} + q^{K_b}(1+q)^{K_a-K_b-1}\right]^{-1}$. Then, the cooperation rate of the two players becomes

$$P(AoN_{K_a}, AoN_{K_b}) = \sum_{i=0}^{K_b-1} x_i[\varepsilon^2 + \varepsilon(1-\varepsilon)] + \sum_{i=K_b}^{K_a-1} x_i[\varepsilon(1-\varepsilon) + \tfrac{1}{2}(\varepsilon^2 + (1-\varepsilon)^2)] + x_{K_a}[(1-\varepsilon)^2 + \varepsilon(1-\varepsilon)].$$

Finally, the payoffs of player AoN_{K_a} and player AoN_{K_b} become

$$\begin{aligned}\pi_{AoN_{K_a}} = &\sum_{i=0}^{K_b-1} x_i[\varepsilon^2R + (1-\varepsilon)^2P + \varepsilon(1-\varepsilon)(T+S)] + \sum_{i=K_b}^{K_a-1} x_i[(1-\varepsilon)^2T + \varepsilon^2S + \varepsilon(1-\varepsilon)(P+R)]\\ &+ x_{K_a}[(1-\varepsilon)^2R + \varepsilon^2P + \varepsilon(1-\varepsilon)(T+S)]\end{aligned}$$

and

$$\begin{aligned}\pi_{AoN_{K_b}} = &\sum_{i=0}^{K_b-1} x_i[\varepsilon^2R + (1-\varepsilon)^2P + \varepsilon(1-\varepsilon)(T+S)] + \sum_{i=K_b}^{K_a-1} x_i[(1-\varepsilon)^2S + \varepsilon^2T + \varepsilon(1-\varepsilon)(P+R)]\\ &+ x_{K_a}[(1-\varepsilon)^2R + \varepsilon^2P + \varepsilon(1-\varepsilon)(T+S)],\end{aligned}$$

respectively. The proof is completed by taking x_i into the equations.

Proposition 3 provides a theoretical analysis of the interactions between AoN_K with different thresholds K. It shows that mutual cooperation is difficult to achieve through coordination among these strategies. AoN_K enforces strict reciprocity among players with the same threshold, making it resistant to exploitation by other strategies. If a coplayer attempts to exploit by defecting, AoN_K responds by reducing that player's payoff, thereby deterring further exploitation.

For two players with different thresholds, K_a and K_b ($K_a > K_b$), AoN_{K_b} will reach its threshold and begin cooperating faster than AoN_{K_a}. However, because AoN_{K_a} has not yet reached its threshold, it defects. This mismatch in behavior resets the state of both players to A_0. In this round of interaction, AoN_{K_b} is vulnerable to

exploitation by AoN_{K_a}, which does not imply that simply increasing the threshold always leads to better outcomes. As Proposition 1 indicates, increasing K reduces the effectiveness of reciprocity among the same strategies when errors are present. An excessively large threshold will effectively turn AoN_K into an unconditional defection strategy. Additionally, when both K_a and K_b increase, the relative disadvantage of AoN_{K_b} becomes less significant, and the gap between the two strategies narrows.

The analysis suggests that the threshold K of AoN_K must be neither too large nor too small; only a moderate threshold can ensure success. Similarly, *AdaCo* shares a strict cooperation condition with AoN_K, requiring the threshold to be reached before cooperation can occur. However, as shown in Proposition 2, *AdaCo* enhances the cooperative potential of AoN_K by maintaining a higher level of cooperation at larger thresholds. This further improves the overall effectiveness of cooperation, particularly as the threshold increases.

Next, we explore the competition of AoN_K and *AdaCo* against two unconditional strategies ($ALLC$ and $ALLD$). Similar to the proof of Proposition 1, we have the following.

Proposition 4. Consider an infinitely repeated prisoner's dilemma with error rate $\varepsilon > 0$ and let $q = \varepsilon^2 + (1-\varepsilon)^2$.

- If one player uses AoN_K and the other player uses $ALLC$, their payoffs are given by

$$\begin{aligned}\pi(AoN_K, ALLC) = \frac{1}{1-(1-2q)(1-q)^{K-1}}\Big[&\left[1-(1-q)^K\right]\left((1-\varepsilon)^2T + \varepsilon(1-\varepsilon)(R+P) + \varepsilon^2 S\right)\\ &+ q(1-q)^{K-1}\left((1-\varepsilon)^2R + \varepsilon(1-\varepsilon)(T+S) + \varepsilon^2 P\right)\Big]\end{aligned} \tag{9}$$

and

$$\begin{aligned}\pi(ALLC, AoN_K) = \frac{1}{1-(1-2q)(1-q)^{K-1}}\Big[&\left[1-(1-q)^K\right]\left(\varepsilon^2T + \varepsilon(1-\varepsilon)(R+P) + (1-\varepsilon)^2 S\right)\\ &+ q(1-q)^{K-1}\left((1-\varepsilon)^2R + \varepsilon(1-\varepsilon)(T+S) + \varepsilon^2 P\right)\Big].\end{aligned} \tag{10}$$

- If one player uses AoN_K and the other player uses $ALLD$, their payoffs are given by

$$\begin{aligned}\pi(AoN_K, ALLD) = \frac{1}{1+q^{K-1}-2q^K}\Big[&\left((1-q^K)(1-\varepsilon) + (1-q)q^{K-1}\varepsilon\right)\left(\varepsilon T + (1-\varepsilon)P\right)\\ &+ \left((1-q^K)\varepsilon + (1-q)q^{K-1}(1-\varepsilon)\right)\left(\varepsilon R + (1-\varepsilon)S\right)\Big]\end{aligned} \tag{11}$$

and

$$\begin{aligned}\pi(ALLD, AoN_K) = \frac{1}{1+q^{K-1}-2q^K}\Big[&(1-\varepsilon)\big[(1-q^K)\varepsilon + (1-q)q^{K-1}(1-\varepsilon)\big]T + \varepsilon\big[(1-q^K)\varepsilon + (1-q)q^{K-1}(1-\varepsilon)\big]R\\ &+ \big[(1-q^K(1-\varepsilon)^2) + (1-q)q^{K-1}\varepsilon(1-\varepsilon)\big]P + \varepsilon\big[(1-q^K)(1-\varepsilon) + (1-q)q^{K-1}\varepsilon\big]S\Big].\end{aligned} \tag{12}$$

As for *AdaCo* with tolerance t, we have the following proposition.

Proposition 5. Consider an infinitely repeated prisoner's dilemma with error rate $\varepsilon > 0$ and let $q = \varepsilon^2 + (1-\varepsilon)^2$.

- If one player uses *AdaCo* with parameters K and t, and the other player uses $ALLC$, their payoffs are given by

$$\begin{aligned}\pi(AdaCo, ALLC) = x_0\Big[&\frac{1-(1-q)^K}{q}\left[(1-\varepsilon)^2T + \varepsilon(1-\varepsilon)(R+P) + \varepsilon^2 S\right]\\ &+ \frac{(1-q)^{K-1}}{1-q^t}\left[(1-\varepsilon)^2R + \varepsilon(1-\varepsilon)(T+S) + \varepsilon^2 P\right]\Big]\end{aligned} \tag{13}$$

and

$$\begin{aligned}\pi(ALLC, AdaCo) = x_0\Big[&\frac{1-(1-q)^K}{q}\left[\varepsilon^2T + \varepsilon(1-\varepsilon)(R+P) + (1-\varepsilon)^2 S\right]\\ &+ \frac{(1-q)^{K-1}}{1-q^t}\left[(1-\varepsilon)^2R + \varepsilon(1-\varepsilon)(T+S) + \varepsilon^2 P\right]\Big],\end{aligned} \tag{14}$$

where $x_0 = \left[\frac{1-(1-q)^K}{q} + \frac{(1-q)^{K-1}}{1-q^t}\right]^{-1}$.

• If one player uses $AdaCo$ with parameters K and t, and the other player uses $ALLD$, their payoffs are given by

$$\begin{aligned}\pi(AdaCo, ALLD) = x_0\Big[&\frac{1-q^K}{1-q}\left[(1-\varepsilon)^2P + \varepsilon(1-\varepsilon)(T+S) + \varepsilon^2R\right] \\ &+ \frac{q^{K-1}}{1-(1-q)^t}\left[(1-\varepsilon)^2S + \varepsilon(1-\varepsilon)(R+P) + \varepsilon^2T\right]\Big]\end{aligned} \tag{15}$$

and

$$\begin{aligned}\pi(ALLD, AdaCo) = x_0\Big[&\frac{1-q^K}{1-q}\left[(1-\varepsilon)^2P + \varepsilon(1-\varepsilon)(T+S) + \varepsilon^2R\right] \\ &+ \frac{q^{K-1}}{1-(1-q)^t}\left[(1-\varepsilon)^2T + \varepsilon(1-\varepsilon)(R+P) + \varepsilon^2S\right]\Big],\end{aligned} \tag{16}$$

where $x_0 = \left[\frac{1-q^K}{1-q} + \frac{q^{K-1}}{1-(1-q)^t}\right]^{-1}$.

Propositions 4 and 5 provide the payoff expressions of AoN_K and $AdaCo$ when interacting with the unconditional benchmarks $ALLC$ and $ALLD$. This analysis is important because these benchmarks reflect two key aspects of whether a memory-based strategy can function as a stable cooperative strategy. First, a successful strategy must resist exploitation by unconditional defection; in other words, $ALLD$ should not gain a significant payoff advantage. Second, the strategy must also be robust to indirect invasion. The well-known example is TFT: it is outperformed by $ALLC$, and $ALLC$ is subsequently invaded by defectors, leading to the eventual breakdown of cooperation. To avoid such indirect collapse, a stable cooperative strategy must be able to exploit $ALLC$, thereby preventing neutral drift toward unconditional cooperation. These propositions show that both AoN_K and $AdaCo$ successfully prevent the indirect collapse driven by $ALLC$ and remain resistant to direct exploitation by $ALLD$, demonstrating stability on both fronts.

Next, let us consider the case where player 1 adopts an AoN_K (or $AdaCo$), while player 2 employs an arbitrary memory-1 strategy denoted by $p = [p_{CC}, p_{CD}, p_{DC}, p_{DD}]$. In this scenario, we examine the infinitely repeated prisoner's dilemma with an error rate of ε. Incorporating implementation error into the strategy, the memory-1 strategy becomes $p' = [p'_{CC}, p'_{CD}, p'_{DC}, p'_{DD}] = (1-\varepsilon)p + \varepsilon(1-p)$. The decision of the memory-1 strategy relies on the outcomes of the most recent round. So, we then add the information of the most recent round to the A_i state, denoted by ${A_0}_a^b$ where player 1's action is a and player 2's action is b in the latest round. In state A_0, the memory-1 strategy needs to specify whether it is ${A_0}_C^D$ or ${A_0}_D^C$. Let ${x_i}_a^b$ represent the probability of the player being in a state ${A_i}_a^b$ $(i = 0, 1, 2, \cdots, K)$ for AoN_K or ${A_i}_a^b$ $(i = 0, 1, 2, \cdots, K+t)$ for $AdaCo$. Here, a and b can take the actions of either C or D. Thus, for the repeated interactions between AoN_K and memory-1 strategies, we have $2K+2$ state probabilities: ${x_0}_C^D$, ${x_0}_D^C$, ${x_1}_C^C$, ${x_1}_D^D$, and so on up to ${x_K}_C^C$ and ${x_K}_D^D$. Since errors are present, all states become ergodic. Similar to the previous analysis, we can establish a system of equations for ${x_i}_a^b$, given by

$$\begin{aligned}{x_0}_C^D = {}&(1-\varepsilon)p'_{DC}\cdot {x_0}_D^C + (1-\varepsilon)p'_{CD}\cdot {x_0}_C^D + (1-\varepsilon)p'_{CC}({x_1}_C^C + {x_2}_C^C + \cdots + {x_{K-1}}_C^C) \\ &+ (1-\varepsilon)p'_{DD}({x_1}_D^D + {x_2}_D^D + \cdots + {x_{K-1}}_D^D) + \varepsilon p'_{CC}({x_K}_C^C + {x_{K+1}}_C^C + \cdots + {x_I}_C^C) \\ &+ \varepsilon p'_{DD}({x_K}_D^D + {x_{K+1}}_D^D + \cdots + {x_I}_D^D),\end{aligned}$$

$$\begin{aligned}{x_1}_C^C &= \varepsilon p'_{DC}\cdot {x_0}_D^C + \varepsilon p'_{CD}\cdot {x_0}_C^D, & {x_1}_D^D &= (1-\varepsilon)(1-p'_{DC}){x_0}_D^C + (1-\varepsilon)(1-p'_{CD}){x_0}_C^D, \\ {x_2}_C^C &= \varepsilon p'_{CC}\cdot {x_1}_C^C + \varepsilon p'_{DD}\cdot {x_1}_D^D, & {x_2}_D^D &= (1-\varepsilon)(1-p'_{CC}){x_1}_C^C + (1-\varepsilon)(1-p'_{DD}){x_1}_D^D, \\ &\vdots & &\vdots \\ {x_{K-1}}_C^C &= (1-\varepsilon)p'_{CC}\cdot {x_{K-2}}_C^C + (1-\varepsilon)p'_{DD}\cdot {x_{K-2}}_D^D & {x_{K-1}}_D^D &= \varepsilon(1-p'_{CC}){x_{K-2}}_C^C + \varepsilon(1-p'_{DD}){x_{K-2}}_D^D, \\ {x_K}_C^C &= (1-\varepsilon)p'_{CC}\cdot {x_{K-1}}_C^C + (1-\varepsilon)p'_{DD}\cdot {x_{K-1}}_D^D + (1-\varepsilon)p'_{CC}\cdot {x_K}_C^C + (1-\varepsilon)p'_{DD}\cdot {x_K}_D^D, \\ {x_K}_D^D &= \varepsilon(1-p'_{CC}){x_{K-1}}_C^C + \varepsilon(1-p'_{DD}){x_{K-1}}_D^D + \varepsilon(1-p'_{CC}){x_K}_C^C + \varepsilon(1-p'_{DD}){x_K}_D^D.\end{aligned}$$

In total, there are $2K+1$ equations. Combined with the equation $\sum_{0\leqslant i\leqslant K}\sum_{a,b}{x_i}_a^b = 1$, we can solve the system of equations to determine the value of ${x_i}_a^b$. After that, we can derive the payoffs of the two players as

$$\begin{aligned}\pi_{AoN_K} &= {x_0}_C^D T + {x_0}_D^C S + \sum_{1\leqslant i\leqslant I}{x_i}_C^C R + \sum_{1\leqslant i\leqslant I}{x_i}_D^D P, \\ \pi_{\text{memory-1}} &= {x_0}_C^D S + {x_0}_D^C T + \sum_{1\leqslant i\leqslant I}{x_i}_C^C R + \sum_{1\leqslant i\leqslant I}{x_i}_D^D P.\end{aligned}$$

The payoff calculation method between $AdaCo$ with cooperation threshold K and tolerance t and the memory-1 strategy is similar to that for AoN_K. We incorporate two state probabilities, ${x_K}_C^D$ and ${x_K}_D^C$, in accordance with the tolerance properties of the strategy. There are $2(K+t)+4$ states. The system of equations is given as follows:

$$\begin{aligned}
{x_0}_C^D &= (1-\varepsilon)p'_{DC}\cdot {x_0}_D^C + (1-\varepsilon)p'_{CD}\cdot {x_0}_C^D + (1-\varepsilon)p'_{CC}({x_1}_C^C + {x_2}_C^C + \cdots + {x_{K-1}}_C^C)\\
&\quad + (1-\varepsilon)p'_{DD}({x_1}_D^D + {x_2}_D^D + \cdots + {x_{K-1}}_D^D) + \varepsilon p'_{CC}({x_K}_C^C + {x_{K+1}}_C^C + \cdots + {x_{K+t-1}}_C^C)\\
&\quad + \varepsilon p'_{DD}({x_K}_D^D + {x_{K+1}}_D^D + \cdots + {x_{K+t-1}}_D^D) + \varepsilon p'_{DC}\cdot {x_K}_D^C + \varepsilon p'_{CD}\cdot {x_K}_C^D,\\
{x_1}_C^C &= \varepsilon p'_{DC}\cdot {x_0}_D^C + \varepsilon p'_{CD}\cdot {x_0}_C^D, \quad {x_1}_D^D = (1-\varepsilon)(1-p'_{DC}){x_0}_D^C + (1-\varepsilon)(1-p'_{CD}){x_0}_C^D,\\
{x_2}_C^C &= \varepsilon p'_{CC}\cdot {x_1}_C^C + \varepsilon p'_{DD}\cdot {x_1}_D^D, \quad {x_2}_D^D = (1-\varepsilon)(1-p'_{CC}){x_1}_C^C + (1-\varepsilon)(1-p'_{DD}){x_1}_D^D,\\
&\vdots \qquad\qquad\qquad\qquad \vdots\\
{x_{K+1}}_C^C &= (1-\varepsilon)p'_{CC}\cdot {x_K}_C^C + (1-\varepsilon)p'_{DD}\cdot {x_K}_D^D + (1-\varepsilon)p'_{DC}\cdot {x_K}_D^C + (1-\varepsilon)p'_{CD}\cdot {x_K}_C^D,\\
{x_{K+1}}_D^D &= \varepsilon(1-p'_{CC})\cdot {x_K}_C^C + \varepsilon(1-p'_{DD})\cdot {x_K}_D^D + \varepsilon(1-p'_{DC})\cdot {x_K}_D^C + \varepsilon(1-p'_{CD})\cdot {x_K}_C^D,\\
{x_{K+2}}_C^C &= (1-\varepsilon)p'_{CC}\cdot {x_{K+1}}_C^C + (1-\varepsilon)p'_{DD}\cdot {x_{K+1}}_D^D, \quad {x_{K+2}}_D^D = \varepsilon(1-p'_{CC}){x_{K+1}}_C^C + \varepsilon(1-p'_{DD}){x_{K+1}}_D^D,\\
&\vdots \qquad\qquad\qquad\qquad \vdots\\
{x_{K+t}}_C^C &= (1-\varepsilon)p'_{CC}\cdot {x_{K+t-1}}_C^C + (1-\varepsilon)p'_{DD}\cdot {x_{K+t-1}}_D^D + (1-\varepsilon)p'_{CC}\cdot {x_{K+t}}_C^C + (1-\varepsilon)p'_{DD}\cdot {x_{K+t}}_D^D,\\
{x_{K+t}}_D^D &= \varepsilon(1-p'_{CC})\cdot {x_{K+t-1}}_C^C + \varepsilon(1-p'_{DD})\cdot {x_{K+t-1}}_D^D + \varepsilon(1-p'_{CC})\cdot {x_{K+t}}_C^C + \varepsilon(1-p'_{DD})\cdot {x_{K+t}}_D^D,\\
{x_K}_D^C &= (1-\varepsilon)(1-p'_{CC})\cdot {x_{K+t}}_C^C + (1-\varepsilon)(1-p'_{DD})\cdot {x_{K+t}}_D^D \quad {x_K}_C^D = \varepsilon p'_{CC}\cdot {x_{K+t}}_C^C + \varepsilon p'_{DD}\cdot {x_{K+t}}_D^D.
\end{aligned}$$

Combined with the equation $\sum_{0\leqslant i\leqslant K+t}\sum_{a,b} {x_i}_a^b = 1$, there are $2(K+t)+4$ equations. We can determine the value of ${x_i}_a^b$ by solving the system of equations. Thus, the players' corresponding payoffs are

$$\begin{aligned}
\pi_{AdaCo} &= ({x_0}_C^D + {x_K}_C^D)T + ({x_0}_D^C + {x_K}_D^C)S + \sum_{1\leqslant i\leqslant K+t} {x_i}_C^C R + \sum_{1\leqslant i\leqslant K+t} {x_i}_D^D P,\\
\pi_{\text{memory-1}} &= ({x_0}_C^D + {x_K}_C^D)S + ({x_0}_D^C + {x_K}_D^C)T + \sum_{1\leqslant i\leqslant K+t} {x_i}_C^C R + \sum_{1\leqslant i\leqslant K+t} {x_i}_D^D P.
\end{aligned}$$

The above framework allows us to explicitly calculate the payoffs between AoN_K (or $AdaCo$) and any memory-1 strategy. This point matters: for many complex strategies, calculating their accurate payoffs is generally difficult, so their performance is often evaluated through computational simulations. Here, the analytical approach allows these payoffs to be obtained both efficiently and accurately, providing a well-founded basis for subsequent evolutionary dynamics analysis and supporting a more precise understanding of strategy evolution. Moreover, the same approach can be extended to evaluate the payoffs of AoN_K and $AdaCo$ against memory-2 or more general memory-n strategies, with corresponding adjustments to the state space and its transition structure. In this work, however, we focus on memory-1 strategies, which are widely studied and commonly regarded as a standard baseline for evaluating cooperative behavior.

4 Evolutionary dynamics

4.1 Evolutionary dynamics of two-strategy competition

While the payoff analysis above provides explicit formulas, it remains difficult to directly grasp the robustness of strategies from such static comparisons. Evolutionary dynamics offer a complementary perspective, both by yielding more intuitive insights and by testing whether strategies can persist under invasion pressure. A fundamental criterion for the success of a cooperative strategy is its ability to withstand the spread of unconditional defectors ($ALLD$). We therefore consider a well-mixed population of infinite size in which the only competing strategies are $AdaCo$ and $ALLD$. Let x denote the frequency of $AdaCo$ players, with the remaining fraction $1-x$ consisting of $ALLD$ players. The strategy frequencies evolve according to replicator dynamics [42], given by $\dot{x} = x(f_{AdaCo} - \bar{f})$, where f_{AdaCo} is the expected payoff of an $AdaCo$ player and $\bar{f}$ is the population average payoff. This system admits two trivial equilibria ($x=0$ and $x=1$) and one internal equilibrium x_* determined by the condition $f_{AdaCo} = f_{ALLD}$.

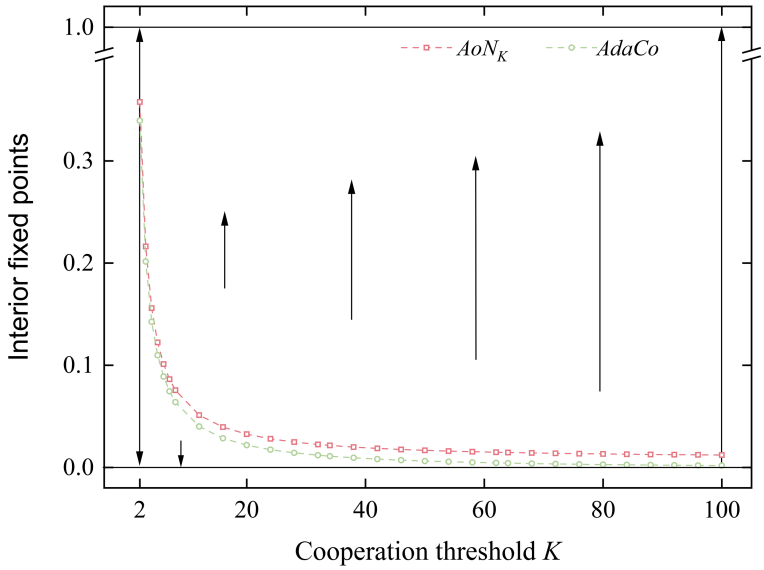

Figure 3 (Color online) Interior fixed point x_* of the replicator dynamics describing competition between $AdaCo$ (or AoN_K) and $ALLD$. We consider an infinite, well-mixed population in which a fraction x adopts $AdaCo$ with $t = 2$ (or AoN_K), and the remaining fraction $1 - x$ plays $ALLD$. The frequency x evolves under the replicator equation $\dot{x} = x\,(f_{AdaCo} - \bar{f})$, where f_{AdaCo} is the expected payoff of an $AdaCo$ player and $\bar{f}$ denotes the mean payoff in the population. A strategy increases in frequency when its payoff exceeds the population mean, and decreases otherwise. For $K \geqslant 2$, this dynamical system exhibits two boundary equilibria ($x = 0$ and $x = 1$) and one internal equilibrium x_*. If the initial frequency x exceeds x_*, the dynamics converge to full dominance of $AdaCo$; if $x < x_*$, $ALLD$ takes over the population. Hence, $1 - x_*$ represents the attraction basin of $AdaCo$, and a larger basin indicates stronger resistance against invasion by $ALLD$. Arrows on the phase line indicate the direction of selection. Similar properties are observed when AoN_K competes with $ALLD$ in the population. Increasing the cooperation threshold K systematically enlarges the attraction basin of $AdaCo$ (or AoN_K), enhancing their robustness against defectors, with the improvement being more pronounced for $AdaCo$.

As illustrated in Figure 3, the population dynamics exhibit positive frequency dependence: if the initial share of $AdaCo$ players exceeds x_*, cooperation spreads until fixation ($x = 1$), whereas otherwise the population converges to full defection ($x = 0$). Similar dynamics is observed when AoN_K competes with unconditional defectors, and increasing K enlarges the basins of attraction for both strategies. Importantly, $AdaCo$ consistently outperforms AoN_K: its basin of attraction is slightly larger, and the cooperative state at $x = 1$ achieves the maximum possible level of cooperation.

We then examine whether $AdaCo$ possesses an evolutionary advantage over other classical strategies in the repeated prisoner's dilemma game. Specifically, let $AdaCo$ compete with each of the widely recognized strategies. To ensure the robustness of our study, we diversify the competition environments by selecting as many classic strategies as possible. The selected strategies include the pure strategies, the stochastic strategies, and complex strategies beyond the description of memory-n strategies. The evolutionary dynamics is described by the replicator dynamics. We found that the internal equilibrium point does not exist for the differential equations describing the pairwise competitions. As a result, the population would eventually be stabilized in a homogeneous state. Figure 4 demonstrates that $AdaCo$ can always pervade into the whole population when in competition with $AllC$, $ALLD$, $GTFT$, $WSLS$, ZD_3 [23], $HardMajority$ [43], and $CumulativeReciprocity$ [25], respectively, for most of the parameter sets. Only for $K = 30$, $AdaCo$ is dominated by the $CumulativeReciprocity$ strategy. As established, classic memory-1 strategies with a certain level of tolerance such as $GTFT$ perform well in many competition settings. However, these strategies are no lack of weaknesses. Let us consider the following simple scenario. The population would neutrally drift when just unconditional cooperators and $GTFT$ compete to survive. Once cooperators dominate, unconditional defectors can expand rapidly. This means that $GTFT$ can be indirectly invaded by defectors. Unlike this indirect evolutionary path, $AdaCo$ can not only resist the invasion of defectors but also win over $GTFT$ directly. These observations illustrate the effectiveness and necessity of combining properties of traditional memory-n strategies in designing strategies in repeated games.

4.2 Evolutionary dynamics of multi-strategy competition

To further verify the effectiveness of $AdaCo$, we consider a larger set of strategies competing in a well-mixed population of finite size M. Let Ω denote the set of strategies available. Following common practice, we use a pairwise comparison process to characterize the population dynamics. The population evolves in discrete time

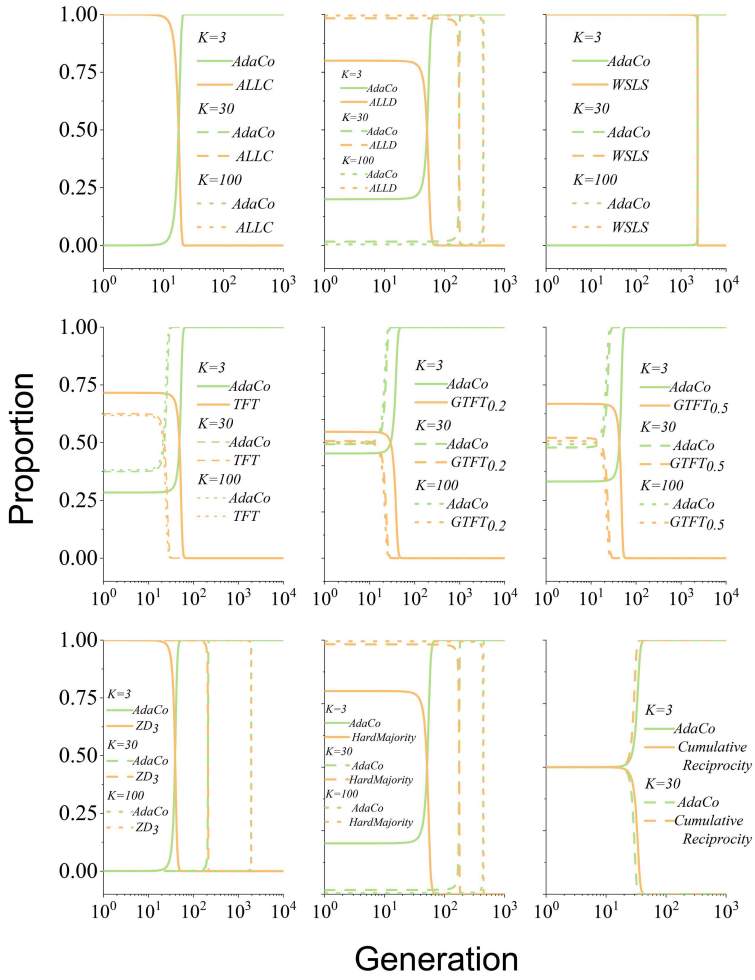

Figure 4 (Color online) Evolutionary dynamics of pairwise competition. We assess the evolutionary performance of $AdaCo$ with three cooperation thresholds, $K = 3$, $K = 30$, and $K = 100$, by allowing it to compete pairwise against a variety of classic strategies in a well-mixed infinite population. Let $x_i(\tau)$ denote the proportion of strategy s_i at generation τ, which evolves according to the discrete replicator rule $x_i(\tau+1) = x_i(\tau)\frac{f(s_i)}{\bar{f}}$, where $f(s_i)$ and $\bar{f}$ are the expected payoff of s_i and the mean population payoff, respectively. Here, $\tau = 1$ indicates the initial composition of the population. Across these competitions, $AdaCo$ requires a sufficiently high initial frequency to dominate in some cases, whereas in others it can expand from rarity. Notably, against $ALLC$, $WSLS$, and ZD_3, all three threshold levels of $AdaCo$ successfully invade even when initially rare. In contrast, when competing with $CumulativeReciprocity$ ($\Delta = 2$), $AdaCo$ ($K = 3$) takes over while $AdaCo(K = 30)$ does not. For the remaining strategies, larger thresholds generally reduce the minimal initial frequency required for $AdaCo$ to prevail, indicating enhanced competitive robustness. Full strategy definitions are provided in the Supplementary Information.

steps. In each evolutionary time step, each player of the population adopts a strategy in Ω. They each interact with all other players and accumulate payoffs. After that, a player (learner) is randomly selected from the population to update the strategy. With probability μ, a mutation happens and the player randomly selects one strategy from the set S. With probability $1-\mu$, learning happens. Then another player is randomly selected from the population as a role model. Let π and $\tilde{\pi}$ denote the payoffs of the learner and the model player, respectively. The learner imitates the strategy of the role model with probability $\frac{1}{1+e^{-\beta(\tilde{\pi}-\pi)}}$. The parameter $\beta > 0$ represents the intensity of selection, weighing the effects of payoffs on evolutionary dynamics. Its essence lies in that strategies yielding higher payoffs are more likely to spread in the population. For small mutation rates, the population would spend most of the time in homogeneous states where just one strategy takes over the whole population. In this situation, the population dynamics can be approximated by an embedded Markov chain consisting of just these homogeneous states. The abundances of strategies correspond to the stable distribution, which can be obtained by solving the eigenvector (associated with the largest eigenvalue) of the embedded Markov chain. For larger mutation rates, we use computer simulations to explore the population dynamics.

We begin by examining the optimal cooperation thresholds for AoN_K when different AoN_K ($K \in \{1, 2, 3, ..., 50\}$) compete in the population. Results show that strategies of varying cooperation thresholds have their respective advantages and weak points. For strategies of low cooperation thresholds, they stabilize the population at very

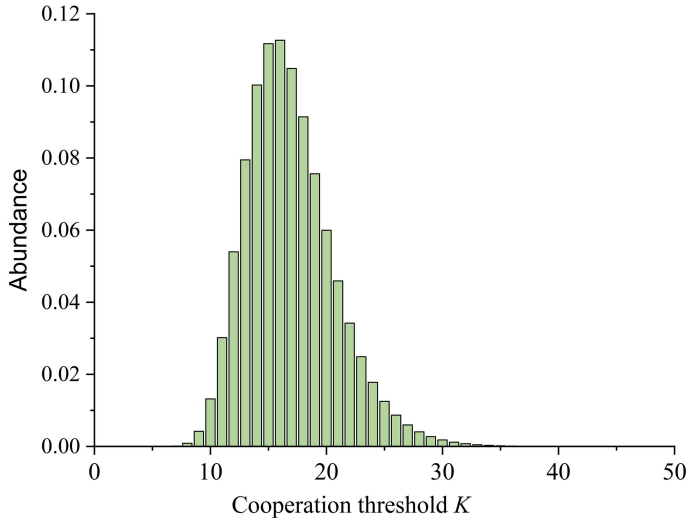

Figure 5 (Color online) Evolutionary dynamics within the class of AoN_K. We examine a family of 50 strategies, AoN_K with $K = 1, 2, \ldots, 50$, competing in a well-mixed population by means of pairwise comparison dynamics. In each update step, a randomly chosen individual adopts the strategy of another with a probability increasing in the payoff difference, while with a small probability μ mutation introduces a randomly selected strategy. Under rare mutation, the population spends most of the time in homogeneous states; consequently, the long-run abundance of each strategy can be computed, which reflects its evolutionary performance. We find that the abundance of AoN_K is maximized at an intermediate cooperation threshold: moving away from this value in either direction leads to reduced prevalence. Strategies with low thresholds maintain cooperation under noise but are vulnerable to invasion by high-threshold strategies; conversely, excessively high thresholds allow exploitation of more cooperative types but fail to sustain mutual cooperation when competing among themselves. Overall, the most abundant strategy in the long run is AoN_{16}. Parameters: population size $M = 100$, implementation error $\varepsilon = 0.01$, selection intensity $\beta = 1$.

high levels of cooperation in the presence of noise but are vulnerable to invasion of strategies of higher cooperation thresholds. Strategies of high cooperation thresholds can still exploit strategies of shorter memories. However, interactions between the strategies with larger K are unable to achieve high levels of cooperation, weakening the stability of the homogeneous populations of AoN_K with larger K. In between comes a moderate cooperation threshold which proves to be the most effective for AoN_K. This optimal memory length may depend on the specific competition environments including the competing strategies, the noise effect, and the game parameters. As the competition environment is considered now, the strategy AoN_{16} enjoys the largest abundance in the long run (see Figure 5).

To further explore the performance and robustness of *AdaCo*, two more complex competition environments are considered (Figure 6). In the first environment, ten classic strategies $GTFT_{0.4}$, $GTFT_{0.2}$, $WSLS$, $ALLD$, $GRIM$, $ALLC$, $RANDOM$, TFT, ZD_2, ZD_4 are selected to compete with *AdaCo*. These strategies are typical as the good performances of some of those strategies have been repeatedly found in respective competition environments [44]. In the absence of *AdaCo*, $GTFT_{0.2}$ accounts for around 46.8% of the population in the long run. $GTFT_{0.4}$ comes in the second place and its abundance is around 20.5%. Though not so abundant, the three strategies $ALLD$, $WSLS$, and $GRIM$ can also take over the population from time to time, and thus their fractions are non-negligible. In addition, extortionate strategies have no chance to evolve, in line with the results reported in [45]. The rationale underlying this observation is that the competitions between these strategies exhibit quasi-cyclical dominance. Each strategy can just be effective in competing with some of the ten strategies. In competition with the remaining strategies, this strategy would not be so effective or even be dominated. Using more history information in decision-making can overcome the limitations of memory-1 strategies. *AdaCo* is such a decision-making rule we have proposed. When *AdaCo* enters the competition, the population exhibits qualitatively different dynamics. Figure 6 shows that in competition with the ten selected strategies, *AdaCo* enjoys an abundance of nearly 100%, indicating that *AdaCo* players can outcompete the selected strategies and take over the population almost all the time. This should be attributed to the integration of advantages in *AdaCo*, of previous strategies. Moreover, the cooperation level is around 83% in the absence of *AdaCo*. With *AdaCo* present, the cooperation level can be improved as high as 100%. This means that classic memory-1 strategies are unable to invade the population occupied by *AdaCo* players.

In another typical competition environment, we considered a large set of memory-1 strategies. Following expression (1), a memory-1 strategy can be represented as a vector $[p_0, p_{CC}, p_{CD}, p_{DC}, p_{DD}]$, where p_0 represents the probability of cooperating in the first round and p_{XY} represents the player's probability of cooperation when the player has chosen $X \in \{C, D\}$ and the opponent has chosen $Y \in \{C, D\}$ in the previous round. Under a noisy environment, interactions between memory-1 strategies would enter stable distributions. Thus p_0 has no effect on the

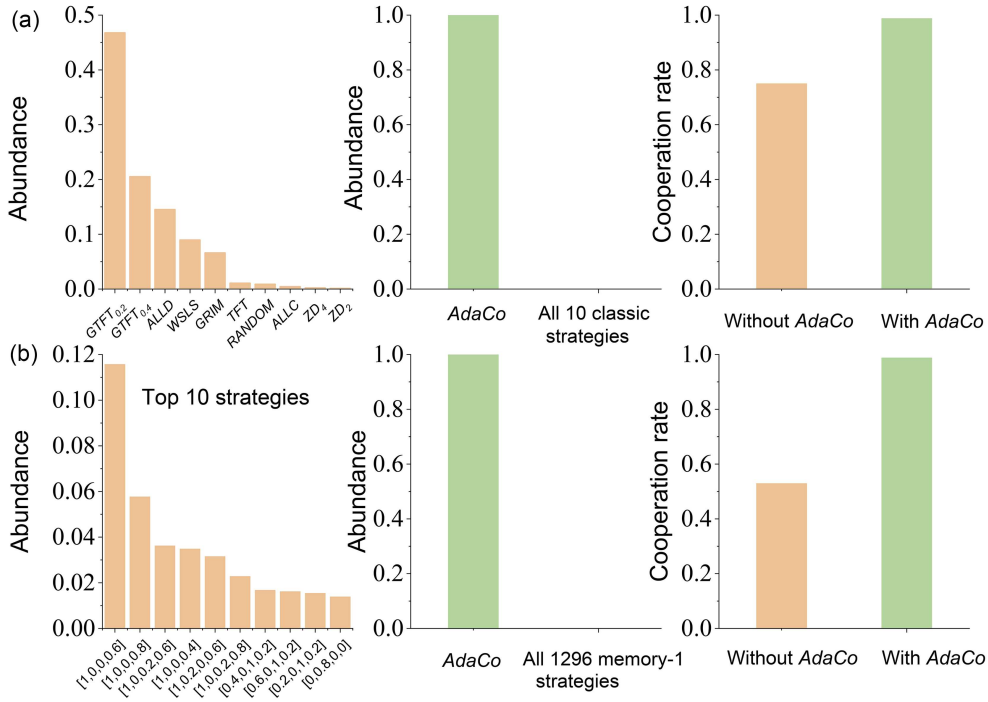

Figure 6 (Color online) *AdaCo*'s evolutionary performance in competition with classic strategies. We evaluate evolutionary outcomes using the standard pairwise comparison process in a well-mixed population of size $M = 100$. At each update step, a randomly chosen individual (the learner) may adopt the strategy of another individual (the role model), with probability $\frac{1}{1+e^{-\beta(\tilde{\pi}-\pi)}}$ depending on the payoff difference between them. Thus, strategies yielding higher payoffs are more likely to spread in the population. Occasional mutations introduce alternative strategies, ensuring that the system explores the strategy set over time. The resulting long-run abundances indicate each strategy's evolutionary success. (a) We first consider *AdaCo* ($K = 3$, $t = 2$) competing with ten widely studied strategies. In the absence of *AdaCo*, the generous strategy $GTFT_{0.2}$ attains the highest abundance, with $GTFT_{0.4}$, $ALLD$, $WSLS$, and $GRIM$ also maintaining non-negligible presence. When *AdaCo* is introduced, it overwhelmingly dominates the population, and the overall cooperation level increases significantly. (b) We then examine competition between *AdaCo* and the full space of memory-1 strategies. Each memory-1 strategy is parameterized as $[P_{CC}, P_{CD}, P_{DC}, P_{DD}]$, and we sample a $6\times6\times6\times6$ grid in $[0,1]^4$, yielding 1296 strategies. Without *AdaCo*, several memory-1 strategies form a quasi-cyclic dominance pattern with comparable abundances. Once *AdaCo* is included, this cycle collapses and *AdaCo* nearly fixes in the population, again raising the overall cooperation rate. Parameters: implementation error $\varepsilon = 0.01$ and selection intensity $\beta = 1$.

long-term payoffs. To simplify our simulations, the values of p_{XY} are constrained to the set $\{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$. Consequently, our simulation encompassed a total of 6^4 memory-1 strategies. As a control, we first let these memory-1 strategies compete in the evolutionary race. Results are shown in Figure 6. Although various strategies have a chance to take over the population, the $WSLS$-like strategies are the most prevalent. Due to the high level of diversity of strategies, even the most abundant strategy $[1, 0, 0, 0.6]$ just enjoys a fraction of 11.5% of the time, further corroborating the ebbing effectiveness of each memory-1 strategy in interacting with diverse memory-1 strategies. We then incorporate *AdaCo* into the competition environment. As mentioned above, *AdaCo* player is tolerant but to a limited degree. This tolerance has a remarkable effect in improving payoffs when interactions happen between *AdaCo* players themselves. Figure 6 shows that *AdaCo* strategy does have a good performance as it can wipe out all other memory-1 strategies and take over the population almost all the time. The abundance of all memory-1 strategies combined is vanishingly small. The abundance of *AdaCo* is above 99%. In the absence of *AdaCo*, the cooperation level in the population is further reduced as the competing strategies diversify. This is understandable. Some strategies can achieve high cooperation levels, while others cannot. The population exhibits quasi-cyclical dominance. On average, the cooperation level in the population is discounted. Once *AdaCo* is introduced to the population, it can wipe out all memory-1 strategies. The quasi-cyclical dominance disappears. The population exhibits a very high level of cooperation, as interactions between *AdaCo* players can achieve mutual reciprocity effectively due to the inherent tolerance of *AdaCo*.

We would like to emphasize that the tolerance in *AdaCo* entails a certain level of risk of being exploited. For instance, players can exploit *AdaCo* players by defecting during a stable cooperative state while cooperating in a non-stable cooperative state. To implement such exploitation, players need an equal or greater amount of history information than *AdaCo* does for decision-making in playing the repeated prisoner's dilemma game. Nevertheless, even if such an exploitation is successful, the increment of cooperation thresholds of *AdaCo* strategy reduces the possibility of being exploited and thus enhances the evolutionary competence of *AdaCo*. In fact, in these competition

scenarios as considered above, no strategies can identify *AdaCo*'s tolerance and exploit it. As a result, the population can always be stabilized at high levels of cooperation when *AdaCo* players take part in evolutionary races.

5 Discussions and conclusion

Understanding the emergence and stability of cooperation in repeated cooperative dilemmas requires a careful examination of how individuals leverage past interactions to guide future actions [46]. Memory-based strategies, where decisions depend on the most recent n rounds, provide a natural framework to formalize this process. Memory-1 strategies, such as TFT and $WSLS$, have been extensively studied due to computational simplicity and their ability to sustain cooperation in small-scale interactions [47, 48]. However, short memory imposes limits on the ability to robustly identify and respond to cooperative partners, motivating the exploration of longer memory strategies.

Among memory-n strategies, the AoN_K class exhibits a particularly effective coordination mechanism: "cooperate when synchronized, otherwise defect". This simple yet powerful rule enables rapid establishment of mutual cooperation when all players follow it, reinforcing stable reciprocity in homogeneous populations. Evolutionary simulations and analytical payoffs reveal that AoN_K excels in short-memory settings, achieving self-cooperation, error correction, and robust retaliation against defectors. Nevertheless, as memory length increases, this coordination rule becomes vulnerable: implementation errors can cascade across multiple rounds, breaking synchronization and undermining cooperation. Consequently, AoN_K with extended memory is no longer guaranteed to maintain high cooperation rates, highlighting that coordination alone is insufficient for long-term stability under realistic noisy environments.

To tackle these limitations, we introduced the *AdaCo* strategy, which preserves the coordination principle of AoN_K while incorporating tolerance. *AdaCo* players track consecutive coordinated rounds and initiate cooperation once a threshold is reached. Critically, after sufficient coordination, they tolerate occasional defections without fully reverting to defection. This design mitigates error propagation while sustaining mutual reciprocity. Analytical derivations and evolutionary simulations demonstrate that *AdaCo* consistently outperforms AoN_K across a wide range of cooperation thresholds, tolerance levels, group sizes, and competition scenarios, including interactions with classic memory-1 strategies such as $ALLC$, $ALLD$, TFT, $WSLS$, $GTFT$, ZD, and recently discovered strategies like *CumulativeReciprocity*.

Our results also highlight the stability of *AdaCo* in population dynamics. In well-mixed populations, *AdaCo* can resist invasions by unconditional defectors ($ALLD$) more effectively than AoN_K, enlarging the basins of attraction for cooperative equilibria and stabilizing high levels of cooperation. Moreover, in heterogeneous populations where multiple strategies coexist, *AdaCo* overwhelmingly dominates, suggesting that the combination of coordination and tolerance provides a general mechanism for sustaining cooperation. These findings reinforce that properly leveraging historical information, not merely increasing memory length, is crucial for strategy design in repeated cooperative dilemmas.

In summary, both AoN_K and *AdaCo* exemplify how simple coordination principles can promote stable cooperation. While AoN_K relies on the immediate feedback of synchronized actions, *AdaCo* extends this framework by integrating tolerance, allowing cooperative behavior to persist under noisy conditions and extended interactions. Our work highlights the importance of balancing coordination with flexibility and provides a foundation for exploring more sophisticated history-based strategies that may underlie the emergence and maintenance of cooperation in complex social and ecological systems.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 62273266, 62533002, 62036002).

Supporting information The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Axelrod R, Hamilton W D. The evolution of cooperation. Science, 1981, 211: 1390–1396
2 Axelrod R, Dion D. The further evolution of cooperation. Science, 1988, 242: 1385–1390
3 Kerr B, Godfrey-Smith P, Feldman M W. What is altruism? Trends Ecol Evol, 2004, 19: 135–140

4 Rand D G, Nowak M A. Human cooperation. Trends Cogn Sci, 2013, 17: 413–425
5 Kleshnina M, Hilbe C, Šimsa Š, et al. The effect of environmental information on evolution of cooperation in stochastic games. Nat Commun, 2023, 14: 4153
6 Van Segbroeck S, Pacheco J M, Lenaerts T, et al. Emergence of fairness in repeated group interactions. Phys Rev Lett, 2012, 108: 158104
7 Wang Y, Li A, Wang L. Networked dynamic systems with higher-order interactions: stability versus complexity. Natl Sci Rev, 2024, 11: nwae103
8 Wang G, Su Q, Wang L, et al. Reproductive variance can drive behavioral dynamics. Proc Natl Acad Sci USA, 2023, 120: e2216218120
9 Su Q, McAvoy A, Mori Y, et al. Evolution of prosocial behaviours in multilayer populations. Nat Hum Behav, 2022, 6: 338–348
10 Braga I, Wardil L. When stochasticity leads to cooperation. Phys Rev E, 2022, 106: 014112
11 Wu T, Wang L, Fu F, et al. Coevolutionary dynamics of phenotypic diversity and contingent cooperation. PLoS Comput Biol, 2017, 13: e1005363
12 Fu F, Hauert C, Nowak M A, et al. Reputation-based partner choice promotes cooperation in social networks. Phys Rev E, 2008, 78: 026117
13 Sheng A, Su Q, Wang L, et al. Strategy evolution on higher-order networks. Nat Comput Sci, 2024, 4: 274–284
14 Nowak M, Sigmund K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. Nature, 1993, 364: 56–58
15 Chen X, Wang L. Promotion of cooperation induced by appropriate payoff aspirations in a small-world networked game. Phys Rev E, 2008, 77: 017103
16 Fischer I, Frid A, Goerg S J, et al. Fusing enacted and expected mimicry generates a winning strategy that promotes the evolution of cooperation. Proc Natl Acad Sci USA, 2013, 110: 10229–10233
17 Hauert C, Stenull O. Simple adaptive strategy wins the prisoner's dilemma. J Theor Biol, 2002, 218: 261–272
18 Hauser O P, Hilbe C, Chatterjee K, et al. Social dilemmas among unequals. Nature, 2019, 572: 524–527
19 McAvoy A, Allen B, Nowak M A. Social goods dilemmas in heterogeneous societies. Nat Hum Behav, 2020, 4: 819–831
20 Zhou L, Wu B, Du J, et al. Aspiration dynamics generate robust predictions in heterogeneous populations. Nat Commun, 2021, 12: 3250
21 Feng X, Zhang Y, Wang L. Evolution of stinginess and generosity in finite populations. J Theor Biol, 2017, 421: 71–80
22 Chen F, Zhou L, Wang L. Cooperation among unequal players with aspiration-driven learning. J R Soc Interface, 2024, 21: 20230723
23 Press W H, Dyson F J. Iterated prisoners dilemma contains strategies that dominate any evolutionary opponent. Proc Natl Acad Sci USA, 2012, 109: 10409–10413
24 Trivers R L. The evolution of reciprocal altruism. Q Rev Biol, 1971, 46: 35–57
25 Li J, Zhao X, Li B, et al. Evolution of cooperation through cumulative reciprocity. Nat Comput Sci, 2022, 2: 677–686
26 Mathieu P, Delahaye J P. New winning strategies for the iterated prisoner's dilemma. JASSS, 2017, 20: 12
27 Li J, Kendall G. The effect of memory size on the evolutionary stability of strategies in iterated prisoner's dilemma. IEEE Trans Evol Computat, 2014, 18: 819–826
28 Imhof L A, Fudenberg D, Nowak M A. Tit-for-tat or win-stay, lose-shift? J Theor Biol, 2007, 247: 574–580
29 Matsen F A, Nowak M A. Win C stay, lose C shift in language learning from peers. Proc Natl Acad Sci USA, 2004, 101: 18053–18057
30 Kim M, Choi J-K, Baek S K. Win-Stay-Lose-Shift as a self-confirming equilibrium in the iterated prisoners dilemma. Proc R Soc B, 2021, 288: 20211021
31 Pinheiro F L, Vasconcelos V V, Santos F C, et al. Evolution of all-or-none strategies in repeated public goods dilemmas. PLoS Comput Biol, 2014, 10: e1003945
32 Hilbe C, Martinez-Vaquero L A, Chatterjee K, et al. Memory-n strategies of direct reciprocity. Proc Natl Acad Sci USA, 2017, 114: 4715–4720
33 Hauert C, Schuster H G. Effects of increasing the number of players and memory size in the iterated prisoner's dilemma: a numerical approach. Proc R Soc Lond B, 1997, 264: 513–519
34 Stewart A J, Plotkin J B. Small groups and long memories promote cooperation. Sci Rep, 2016, 6: 26889
35 Wu T, Fu F, Wang L. Coevolutionary dynamics of aspiration and strategy in spatial repeated public goods games. New J Phys, 2018, 20: 063007
36 Nowak M A. Five rules for the evolution of cooperation. Science, 2006, 314: 1560–1563
37 Szolnoki A, Chen X. Cooperation and competition between pair and multi-player social games in spatial populations. Sci Rep, 2021, 11: 12101
38 Zhang F, Wang X, Wu T, et al. Evolutionary dynamics under partner preferences. J Theor Biol, 2023, 557: 111340
39 Wu T, Fu F, Wang L. Evolutionary games and spatial periodicity. J Autom Intell, 2023, 2: 79–86
40 Hauert C, De Monte S, Hofbauer J, et al. Volunteering as red queen mechanism for cooperation in public goods games. Science, 2002, 296: 1129–1132
41 Wang J, Fu F, Wu T, et al. Emergence of social cooperation in threshold public goods games with collective risk. Phys Rev E, 2009, 80: 016101
42 Cressman R, Tao Y. The replicator equation and other game dynamics. Proc Natl Acad Sci USA, 2014, 111: 10810–10817
43 Mittal S, Deb K. Optimal strategies of the iterated prisoner's dilemma problem for multiple conflicting objectives. IEEE Trans Evol Computat, 2009, 13: 554–565
44 Chen X, Fu F. Outlearning extortioners: unbending strategies can foster reciprocal fairness and cooperation. PNAS Nexus, 2023, 2: 1–12
45 Stewart A J, Plotkin J B. From extortion to generosity, evolution in the iterated prisoners dilemma. Proc Natl Acad Sci USA, 2013, 110: 15348–15353
46 Imhof L A, Nowak M A. Stochastic evolutionary dynamics of direct reciprocity. Proc R Soc B, 2010, 277: 463–468
47 Glynatsi N E, Knight V A. Using a theory of mind to find best responses to memory-one strategies. Sci Rep, 2020, 10: 17287
48 Sheng A, Li A, Wang L, et al. Evolutionary dynamics on sequential temporal networks. PLoS Comput Biol, 2023, 19: e1011333