SCIENCE CHINA
Information Sciences

• Supplementary File •

# 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

In the following, we present a complete description of the mathematical and numerical methods employed to obtain the outcomes presented in the main body of the text. Firstly, we present a mathematical analysis of the cooperation rate and payoff for the $AoN_K$ strategy and $AdaCo$ strategy in the context of self-play, for arbitrary cooperation thresholds $K$. Assessing a strategy's performance through self-cooperative maintenance is crucial, particularly in the presence of error perturbations, as it signifies the ability of a cooperative strategy to effectively sustain cooperation. Furthermore, it is noteworthy to examine how this strategy performs against a variety of other strategies. On the one hand, we derive the expression for the payoff in the game between any two distinct $AoN_K$ strategies. on the other hand, We obtain the payoff of the $AoN_K$ (or $AdaCo$ with cooperation threshold $K$ and tolerance $t$) strategy versus any of the memory-1 strategies by solving a $2K+2$ (or $2(K+t)+4$)-dimensional system of linear equations. Additionally, numerical simulations are utilized to determine the game payoffs for other strategies mentioned in the paper. Finally, we provide a detailed description of the evolutionary dynamics within the populations described in the main body of the paper. This description elucidates how the strategies evolve and interact with each other in the population over time, shedding light on the dynamics and outcomes observed in the study.

## 1 Mathematical analysis of $AoN_K$ and $AdaCo$ strategies for arbitrary memory

Consider an infinitely repeated social interaction in a group of $N$ players. Players decide simultaneously each round to cooperate (C) or defect (D). $AoN_K$ players will cooperate if all members behaved the same in the past $K$ rounds, otherwise, they will defect. So, $K$ is the minimum memory length required by the strategy. $AoN_K$ is very sensitive to error when $K$ is large. To improve the cooperative robustness of $AON_K$ under long memory, we propose the $AdaCo$ strategy. It is reasonable to be tolerant of or ignore occasional errors when players have cooperated for a certain number of consecutive rounds. If player behavior is uncoordinated in the current round, instead of immediately retaliating for round $K$ as $AoN_K$ does, $AdaCo$ will tolerate once (continue to cooperate) if all players have cooperated in the past $t$ rounds. $t$ is the tolerance of the strategy, and the smaller it is, the more tolerant it is.

**Proposition 1.** [$AoN_K$ vs. $AoN_K$] Consider an infinitely repeated social dilemma in a group of N players. Let $\varepsilon > 0$ represent the probability of a player making an error in any given round. Suppose all $N$ players use the $AoN_K$ strategy. Then, the group's average cooperation rate is

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

In the case of the prisoner's dilemma with possible payoffs $R$, $S$, $T$, and $P$, the payoff for each player can be expressed as:

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

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

*Proof.* We categorize the outcomes of the past $K$ steps into $K+1$ states based on the number of latest consistent rounds, denoted as $A_0, A_1, A_2, \cdots, A_K$. $A_i$ means that all players chose the same action in the most recent $i$ rounds. The state $A_0$ means that the most recent round of players' actions was inconsistent. We only focus on the last $K$ steps so the maximum value of $i$ is $K$. $x_i$ denotes the probability that the player is in state $A_i$ in the repeated game. Let $q = \varepsilon^N + (1-\varepsilon)^N$, then we get

$$\begin{aligned} x_i &= qx_{i-1}, i = 1, 2, \cdots, K-1, \\ x_K &= qx_{K-1} + qx_K. \end{aligned}$$

Since all probabilities must end up being one, so

$$\sum_{i=0}^{K} x_i = 1.$$

We get

$$\begin{aligned} x_i &= (1-q)q^{i-1}, i = 1, 2, \cdots, K-1, \\ x_K &= q^K. \end{aligned}$$

Now, let's calculate the average cooperation rate of the 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 $K$ to infinity, 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

$$\begin{aligned} \rho_{AoN_K} =& \sum_{i=0}^{K-1} x_i \left[ C_N^N\varepsilon^N + \frac{N-1}{N}C_N^{N-1}\varepsilon^{N-1}(1-\varepsilon) + \cdots + 0 \cdot C_N^0(1-\varepsilon)^N \right] \\ &+ x_K \left[ C_N^N(1-\varepsilon)^N + \frac{N-1}{N}C_N^{N-1}(1-\varepsilon)^{N-1}\varepsilon + \cdots + 0 \cdot C_N^0\varepsilon^N \right] \\ =& x_0\frac{1-q^K}{1-q}\varepsilon + x_K(1-\varepsilon) \\ =& \varepsilon + (1-2\varepsilon)\left[(1-\varepsilon)^N + \varepsilon^N\right]^K \end{aligned}$$

We use an identical approach to determine player payoffs in the prisoner's dilemma, resulting in the equation

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

**Proposition 2.** [*AdaCo vs. AdaCo*] Consider the same game parameters as in Proposition 1. Suppose all N players use the *AdaCo* strategy with cooperation threshold $K$ and tolerance $t$. The average cooperation rate of the group is

$$\rho_{AdaCo} = x_0 \left[ \frac{1-q^K}{1-q}\varepsilon + \frac{q^K}{(1-q^t)(1-q)}(1-\varepsilon) \right], \tag{3}$$

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

$$\begin{aligned} \pi\left(AdaCo, AdaCo\right) = x_0 \Bigg\{ & \left[ \frac{1-q^K}{1-q}\varepsilon^2 + \frac{q^K}{(1-q^t)(1-q)}(1-\varepsilon)^2 \right] R \\ &+ \left[ \frac{1-q^K}{1-q}(1-\varepsilon)^2 + \frac{q^K}{(1-q^t)(1-q)}\varepsilon^2 \right] P \\ &+ \varepsilon(1-\varepsilon)(T+S) \Bigg\}. \end{aligned} \tag{4}$$

*Proof.* On the basis of Assumption 1, we add $t$ states $(A_{K,1}, A_{K,2}, \cdots, A_{K,t})$. We use $x_{K+i}$ $(i = 1, 2, \cdots, t)$ to denote the probability that a player is in state $A_{K,i}$. According to the *AdaCo* rules, the probability $x_i$ satisfies the following equation

$$x_i = q \cdot x_{i-1}, i = 0, 1, \cdots, K$$

$$\begin{aligned} x_i &= q \cdot x_{i-1}, i = K+1, , K+2, \cdots, K+t-1, \\ x_K &= q \cdot x_{K-1} + (1-q)x_{K+t}, \\ x_{K+t} &= qx_{K+t-1} + qx_{K+t}. \end{aligned}$$

Because the probabilities sum to one,

$$1 = \sum_{i=0}^{K+t} x_i = x_0 \frac{1-q^K}{1-q} + \frac{x_K}{1-q}.$$

By solving the above equations, we get

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

Using these probabilities, we can calculate the player's average cooperation rate as

$$\begin{aligned} \rho_{AdaCo} =& \sum_{i=0}^{K-1} x_i \left[ C_N^N \varepsilon^N + \frac{N-1}{N} C_N^{N-1} \varepsilon^{N-1}(1-\varepsilon) + \cdots + 0 \cdot C_N^0 (1-\varepsilon)^N \right] \\ &+ \sum_{i=K}^{K+t} x_i \left[ C_N^N (1-\varepsilon)^N + \frac{N-1}{N} C_N^{N-1} (1-\varepsilon)^{N-1} \varepsilon + \cdots + 0 \cdot C_N^0 \varepsilon^N \right] \\ =& x_0 \frac{1-q^K}{1-q} \varepsilon + x_K \frac{1-\varepsilon}{1-q} \\ =& x_0 \left[ \frac{1-q^K}{1-q} \varepsilon + \frac{q^K}{(1-q^t)(1-q)} (1-\varepsilon) \right]. \end{aligned}$$

When it comes to the two-person prisoner's dilemma, the player's payoff is

$$\begin{aligned} \pi\left(AdaCo, AdaCo\right) =& \sum_{i=0}^{K-1} x_i \left[ (1-\varepsilon)^2 P + \varepsilon^2 R + \varepsilon(1-\varepsilon)(T+S) \right] \\ &+ \sum_{i=K}^{K+t} x_i \left[ (1-\varepsilon)^2 R + \varepsilon^2 P + \varepsilon(1-\varepsilon)(T+S) \right] \\ &= x_0 \left\{ \left[ \frac{1-q^K}{1-q} \varepsilon^2 + \frac{q^K}{(1-q^t)(1-q)} (1-\varepsilon)^2 \right] \cdot R \right. \\ &+ \left[ \frac{1-q^K}{1-q} (1-\varepsilon)^2 + \frac{q^K}{(1-q^t)(1-q)} \varepsilon^2 ] \cdot \right] P \\ &\left. + \varepsilon(1-\varepsilon)(T+S) \right\}. \end{aligned}$$

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

**Proposition 3.** $[AoN_{Ka}\ vs\ AoN_{Kb}]$ Consider the same game parameters as in Proposition 1. Suppose that two players use different $AoN_K$ strategies, denoted as $AoN_{Ka}$ and $AoN_{Kb}$ respectively, and that $Ka > kb$. Then the average cooperation rate is

$$\begin{aligned} \rho(AoN_{Ka}, AoN_{Kb}) = \varepsilon x_0 \frac{1-q^{Kb}}{1-q} + \frac{1}{2} x_0 q^{Kb} [1 - (1-q)^{Ka-Kb}] \\ + (1-\varepsilon) x_0 q^{Kb} (1-q)^{Ka-Kb-1}, \end{aligned} \tag{5}$$

where $q = \varepsilon^2 + (1-\varepsilon)^2$ and $x_0 = \left[ \frac{1-q^{Kb}}{1-q} + q^{Kb-1} - q^{Kb-1}(1-q)^{Ka-Kb} + q^{Kb}(1+q)^{Ka-Kb-1} \right]^{-1}$. And the corre-

sponding payoffs are

$$\begin{aligned}\pi_{AoN_{Ka}} = &\left[\frac{1-q^{Kb}}{1-q}\varepsilon^2 + (q^{Kb-1} - q^{Kb-1}(1-q)^{Ka-Kb})\varepsilon(1-\varepsilon) + q^{Kb}(1-q)^{Ka-Kb-1}(1-\varepsilon)^2\right]x_0R\\ &+\left[\frac{1-q^{Kb}}{1-q}\varepsilon(1-\varepsilon) + (q^{Kb-1} - q^{Kb-1}(1-q)^{Ka-Kb})(1-\varepsilon)^2 + q^{Kb}(1-q)^{Ka-Kb-1}\varepsilon(1-\varepsilon)\right]x_0T\\ &+\left[\frac{1-q^{Kb}}{1-q}\varepsilon(1-\varepsilon) + (q^{Kb-1} - q^{Kb-1}(1-q)^{Ka-Kb})\varepsilon^2 + q^{Kb}(1-q)^{Ka-Kb-1}\varepsilon(1-\varepsilon)\right]x_0S\\ &+\left[\frac{1-q^{Kb}}{1-q}(1-\varepsilon)^2 + (q^{Kb-1} - q^{Kb-1}(1-q)^{Ka-Kb})\varepsilon(1-\varepsilon) + q^{Kb}(1-q)^{Ka-Kb-1}\varepsilon^2\right]x_0P\end{aligned} \tag{6}$$

and

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

respectively.

*Proof.* Since $Ka > Kb$, the minimum number of states that completely describe the process of the game between the strategies is $Ka+1$. We first need to calculate the value of $x_i$. According to the definition of $AoN_K$, we get

$$x_i = \begin{cases} qx_{i-1} & i = 1, \cdots, K_b \\ (1-q)x_{i-1} & i = K_b+1, \cdots, K_a-1 \\ (1-q)x_{i-1} + qx_i & i = K_a. \end{cases}$$

Because the sum of all $x_i$ is one, we get $x_0 = \left[\frac{1-q^{kb}}{1-q} + q^{kb-1} - q^{kb-1}(1-q)^{ka-kb} + q^{kb}(1+q)^{ka-kb-1}\right]^{-1}$. Then, the cooperation rate of the two players becomes

$$\begin{aligned}P(AoN_{Ka}, AoN_{Kb}) = &\sum_{i=0}^{Kb-1} x_i[\varepsilon^2 + \varepsilon(1-\varepsilon)] + \sum_{i=Kb}^{Ka-1} x_i[\varepsilon(1-\varepsilon) + \frac{1}{2}(\varepsilon^2 + (1-\varepsilon)^2)]\\ &+x_{Ka}[(1-\varepsilon)^2 + \varepsilon(1-\varepsilon)].\end{aligned}$$

Finally, the payoffs of player $AoN_{Ka}$ and player $AoN_{Kb}$ become

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

and

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

Plugging $x_i$ into the equations, we will get the result shown by the proposition.

| | $AoN_1$ | $AoN_2$ | $AoN_3$ | $AoN_4$ | $AoN_5$ | $AoN_6$ | $AoN_7$ | $AoN_8$ | $AoN_9$ | $AoN_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $AoN_1$ | 2.951 | 1.351 | 0.559 | 0.536 | 0.535 | 0.535 | 0.535 | 0.535 | 0.535 | 0.535 |
| $AoN_2$ | 2.974 | 2.913 | 1.269 | 0.717 | 0.703 | 0.702 | 0.702 | 0.702 | 0.702 | 0.702 |
| $AoN_3$ | 2.985 | 2.475 | 2.876 | 1.219 | 0.797 | 0.786 | 0.786 | 0.786 | 0.786 | 0.786 |
| $AoN_4$ | 2.985 | 2.329 | 2.176 | 2.839 | 1.186 | 0.845 | 0.837 | 0.837 | 0.837 | 0.837 |
| $AoN_5$ | 2.985 | 2.325 | 1.997 | 1.976 | 2.803 | 1.163 | 0.877 | 0.871 | 0.870 | 0.870 |
| $AoN_6$ | 2.985 | 2.325 | 1.993 | 1.798 | 1.833 | 2.768 | 1.145 | 0.900 | 0.895 | 0.895 |
| $AoN_7$ | 2.985 | 2.325 | 1.993 | 1.793 | 1.664 | 1.726 | 2.734 | 1.131 | 0.917 | 0.913 |
| $AoN_8$ | 2.985 | 2.325 | 1.993 | 1.793 | 1.660 | 1.568 | 1.642 | 2.700 | 1.120 | 0.931 |
| $AoN_9$ | 2.985 | 2.325 | 1.993 | 1.793 | 1.660 | 1.565 | 1.496 | 1.575 | 2.667 | 1.111 |
| $AoN_{10}$ | 2.985 | 2.325 | 1.993 | 1.793 | 1.660 | 1.565 | 1.493 | 1.441 | 1.521 | 2.635 |

**Table 1** The payoff matrix of the game between different $AoN_K$ strategies.

## 2 Competition with unconditional strategies

Next, we explore competition against two unconditional strategies ($ALLC$ and $ALLD$).

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

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

$$\begin{aligned}\pi(AoN_K, ALLC) = &\left[\frac{1-(1-q)^K}{1-(1-2q)(1-q)^{K-1}}(1-\varepsilon)^2 + \frac{q(1-q)^{K-1}}{1-(1-2q)(1-q)^{K-1}}\varepsilon(1-\varepsilon)\right]T\\ &+\left[\frac{1-(1-q)^K}{1-(1-2q)(1-q)^{K-1}}\varepsilon(1-\varepsilon) + \frac{q(1-q)^{K-1}}{1-(1-2q)(1-q)^{K-1}}(1-\varepsilon)^2\right]R\\ &+\left[\frac{1-(1-q)^K}{1-(1-2q)(1-q)^{K-1}}\varepsilon(1-\varepsilon) + \frac{q(1-q)^{K-1}}{1-(1-2q)(1-q)^{K-1}}\varepsilon^2\right]P\\ &+\left[\frac{1-(1-q)^K}{1-(1-2q)(1-q)^{K-1}}\varepsilon^2 + \frac{q(1-q)^{K-1}}{1-(1-2q)(1-q)^{K-1}}\varepsilon(1-\varepsilon)\right]S\end{aligned} \tag{8}$$

and

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

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

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

and

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

*Proof.* The proof is similar to that of Proposition 1.

As for *AdaCo* strategies 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 the *AdaCo* strategy with cooperation threshold $K$ and tolerance $t$ and the other uses $ALLC$, their payoffs are given by

$$\begin{aligned}\pi(AdaCo, ALLC) = &x_0\left[\frac{1-(1-q)^K}{q}(1-\varepsilon)^2 + \frac{(1-q)^{K-1}}{1-q^t}\varepsilon(1-\varepsilon)\right]T \\ &+x_0\left[\frac{1-(1-q)^K}{q}\varepsilon(1-\varepsilon) + \frac{(1-q)^{K-1}}{1-q^t}(1-\varepsilon)^2\right]R \\ &+x_0\left[\frac{1-(1-q)^K}{q}\varepsilon(1-\varepsilon) + \frac{(1-q)^{K-1}}{1-q^t}\varepsilon^2\right]P \\ &+x_0\left[\frac{1-(1-q)^K}{q}\varepsilon^2 + \frac{(1-q)^{K-1}}{1-q^t}\varepsilon(1-\varepsilon)\right]S\end{aligned} \tag{12}$$

and

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

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* strategy with parameters $K$ and $t$ and the other uses $ALLD$, their payoffs are given by

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

and

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

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

*Proof.* The proof is similar to that of Proposition 1.

## 3 Competition with arbitrary memory-1 strategies

Next, let's consider the case where Player 1 adopts an $AoN_K$ (or $AdaCo$) strategy, while Player 2 employs an arbitrary memory-1 strategy denoted as $p = [p_{CC}, p_{CD}, p_{DC}, p_{DD}]$. In this scenario, we examine the infinitely repeated Prisoner's Dilemma with an error rate of $\varepsilon$. 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 as $\{A_0\}_a^b$ where Player 1's action being $a$ and Player 2's action being $b$ 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 $\{A_i\}_a^b$ $(i = 0, 1, 2, \cdots, K)$ or $\{A_{K,i}\}_a^b$ $(i = K+1, K+2, \cdots, K+t)$. Here, $a$ and $b$ can take the actions of either "C" (cooperate) or "D" (defect). 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

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

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

$$\begin{aligned}{x_K}_C^C &= (1-\varepsilon)p'_{CC}\cdot {x_{K-1}}_C^C + (1-\varepsilon)p'_{DD}\cdot {x_{K-1}}_D^D + (1-\varepsilon)p'_{CC}\cdot {x_K}_C^C + (1-\varepsilon)p'_{DD}\cdot {x_K}_D^D \\ {x_K}_D^D &= \varepsilon(1-p'_{CC}){x_{K-1}}_C^C + \varepsilon(1-p'_{DD}){x_{K-1}}_D^D + \varepsilon(1-p'_{CC}){x_K}_C^C + \varepsilon(1-p'_{DD}){x_K}_D^D\end{aligned}$$

The above yields $2K+1$ equations. Combined with the equation $\sum_{0\leqslant i\leqslant K}\sum_{a,b} x_i^b = 1$, 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

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

respectively.

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

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

We computed the payoff between *AdaCo* and any memory-1 strategy by solving a system of linear equations. In the main text, we selected ten representative classic strategies known for their performance in various competitive environments. These include two unconditional strategies and three pure strategies: $WSLS$ $([1,0,0,1])$, $Grim([1,0,0,0])$, $TFT([1,0,1,0])$, as well as five stochastic strategies: $GTFT_{0.2}([1,0.2,1,0.2])$, $GTFT_{0.4}([1,0.4,1,0.4])$, $ZD_2([8/9,1/2,1/3,0])$, $ZD_3([11/13,1/2,7/26,0])$, $Random([0.5,0.5,0.5,0.5])$.

## 4 Competition with some strategies through computer simulation

Furthermore, in addition to traditional memory-based strategies, we also investigate strategies that go beyond the constraints of finite memory. Without sacrificing generality, we examine two distinct types of strategies: 'HardMajority' and 'Cumulative Reciprocity'. In the 'HardMajority' strategy, a player defects in the initial round and continues to do so unless the opponent cooperates more times than defects. On the other hand, in the 'Cumulative Reciprocity' strategy, a player defects only if the opponent's count of defections surpasses the player's count of defections plus $\Delta$; otherwise, the player cooperates. To analyze the outcomes, we employ computer simulations to derive individual payoffs.

## 5 Evolutionary game dynamics

Nash equilibrium is a static property of games between strategies. Instead, evolutionary game theory provides a classic paradigm for studying population games in which individuals can imitate or learn. More specifically, in an evolutionary system involving many players, players adopt a strategy to interact with other individuals and obtain benefits from the interactions. Successful strategies (with higher payoffs) are more likely to be imitated and diffuse in the population. This paper adopts two evolutionary methods for two-strategy and multi-strategy competition respectively. For the pairwise competition between *AdaCo* (or $AoN_K$) strategy and one other strategy in the population, we consider the large population and pay attention to the evolution process of the strategies. For the evolution of multiple strategies in the population, we focus on the abundance of each strategy in the evolutionary competition.

**Description of population evolution with two strategies**

Consider a large, well-mixed population with two strategies, denoted as strategy $A$ and strategy $B$. $x_A$ represents the frequency of strategy $A$ in the population. So $x_B = 1 - x_A$. The fitness of strategy $A$ is defined as the average payoff of a strategy $A$ player playing against the entire population of players, that is $f(A) = x_A\pi(A, B)+x_B\pi(B, A)$. Similarly, $f(B) = x_A\pi(B, A)+x_B\pi(A, B)$. The average fitness of the population is written as $\bar{f} = x_Af(A)+x_Bf(B)$. Next, players will evolve based on their fitness. The frequency of strategy $A$ in the next generation is determined to be $x'_A = x_Af(A)/\bar{f}$. This basic evolutionary step is repeated until the frequency of each strategy does not change. Either the population is stuck with only one strategy, or both strategies coexist.

**Description of multi-strategy population evolution**

Different from the above, we here consider a finite population of size M. Each pair of players interacts and participates in a repeated game. In each generation, a player is randomly selected from the population. With probability $\mu$ (the mutation has occurred), this player will randomly select a strategy from the strategy set to become his new strategy. With probability $1 - \mu$, the player will compare his payoff $\pi$ with the payoff $\widetilde{\pi}$ of a randomly selected member of the population. Assume that the focal player will switch to the other player's strategy with a probability

$$p = \frac{1}{1 + exp[-\beta(\widetilde{\pi} - \pi)]},$$

where $\beta$ is the strength of selection. It describes the effect of payoff on strategy learning. $p = 0.5$ holds regardless of the difference in payoffs if $\beta = 0$. If $\beta \to \infty$, imitation always occurs despite the small difference in payoffs. This process will be repeated a sufficient number of times.

Here, we follow the framework of Imhof and Nowak and assume mutation is rare. It implies that the mutant either reaches the fixation in the population or goes extinct before the next mutant emerges. Therefore, the population is almost always homogeneous, where members all adopt the same strategy. The fixation probability of the mutant can be calculated to be

$$\rho_M = \left(1 + \sum_{i=1}^{N-1} \prod_{j=1}^{i} \exp\left\{-\beta\left[\pi_M(j) - \pi_R(j)\right]\right\}\right)^{-1}.$$

where $\pi_M(j)$ and $\pi_R(j)$ are the payoffs of the mutation strategy and the resident strategy respectively when there are $j$ mutants in the population. $N$ is the population size. Once the mutant reaches fixation or goes extinct, we introduce a new mutant into the resident population. We will repeat this elementary step sufficiently many times. The population will experience different strategies of occupation in the process, as different mutants invade again and again. However, different strategies occupy different proportions, and the evolution of robust strategies will take more time in this process. It is a Markov process that the population switches from one strategy to another. By calculating the corresponding stationary distribution, we can get the abundance of the strategies in the evolution process, just as described in the main text. However, if the mutation rate is large, this method is not feasible. Instead, we resort to computer simulations, iterating the elementary population updating process with $10^9$ mutant strategies per simulation run. This approach offers a pragmatic approximation for the dynamics.

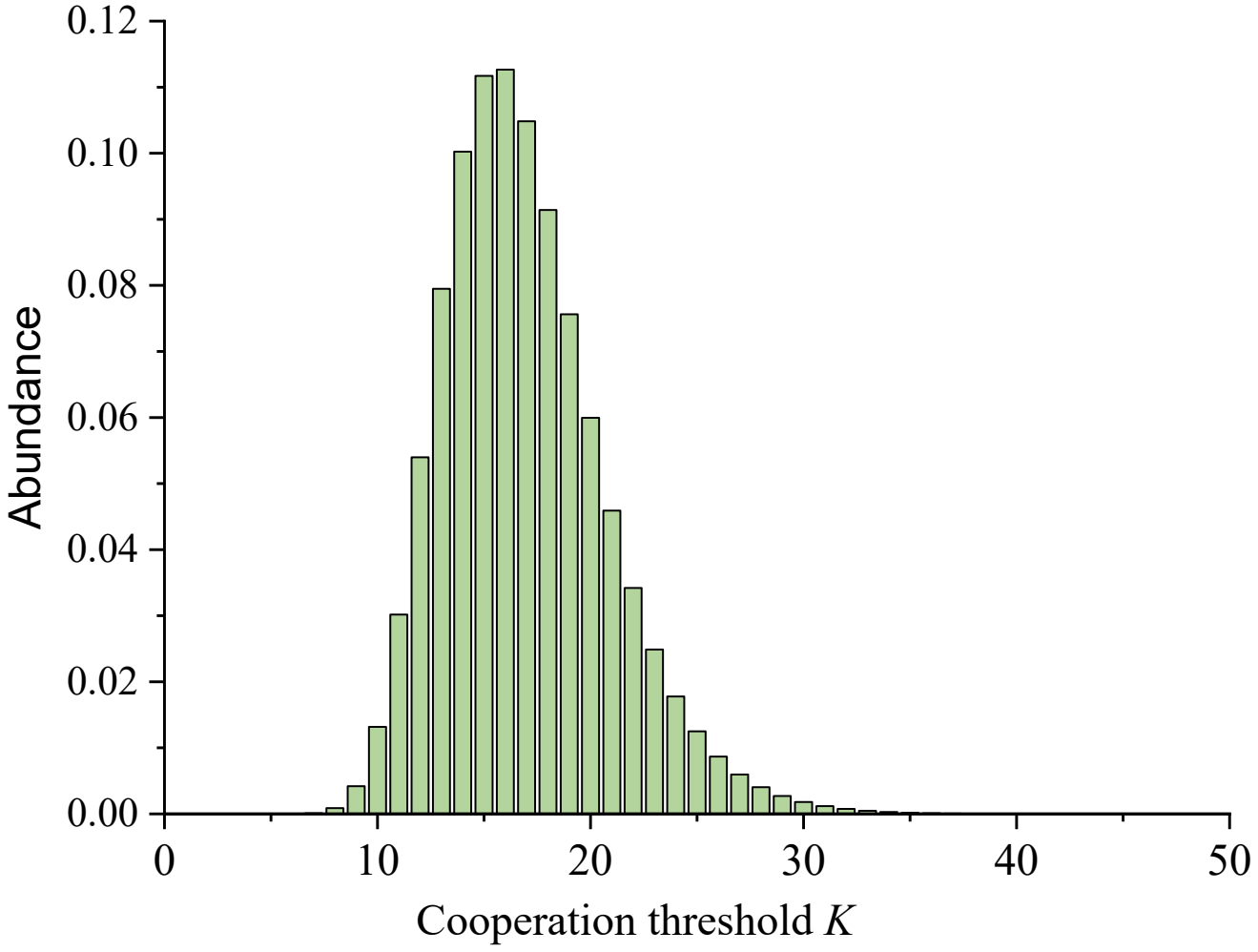


**Figure 1** Evolutionary competition between $AoN_K$ strategy under high mutation rates. We consider a sufficiently large class of 50 strategies ($K = 1, 2, \cdots, 50$) with mutation rates set $\mu = 0.1$. The simulation results show that the $AoN_{17}$ strategy in the middle has the highest abundance, and the strategy abundance on both sides gradually decreases. Similar to the results for rare mutations, the strategy abundance always peaks at a moderate memory length. Other parameters include a population size of $M = 100$, $\varepsilon = 0.01$, and selection strength of $\beta = 1$; The simulation is executed for $10^9$ mutant strategies.

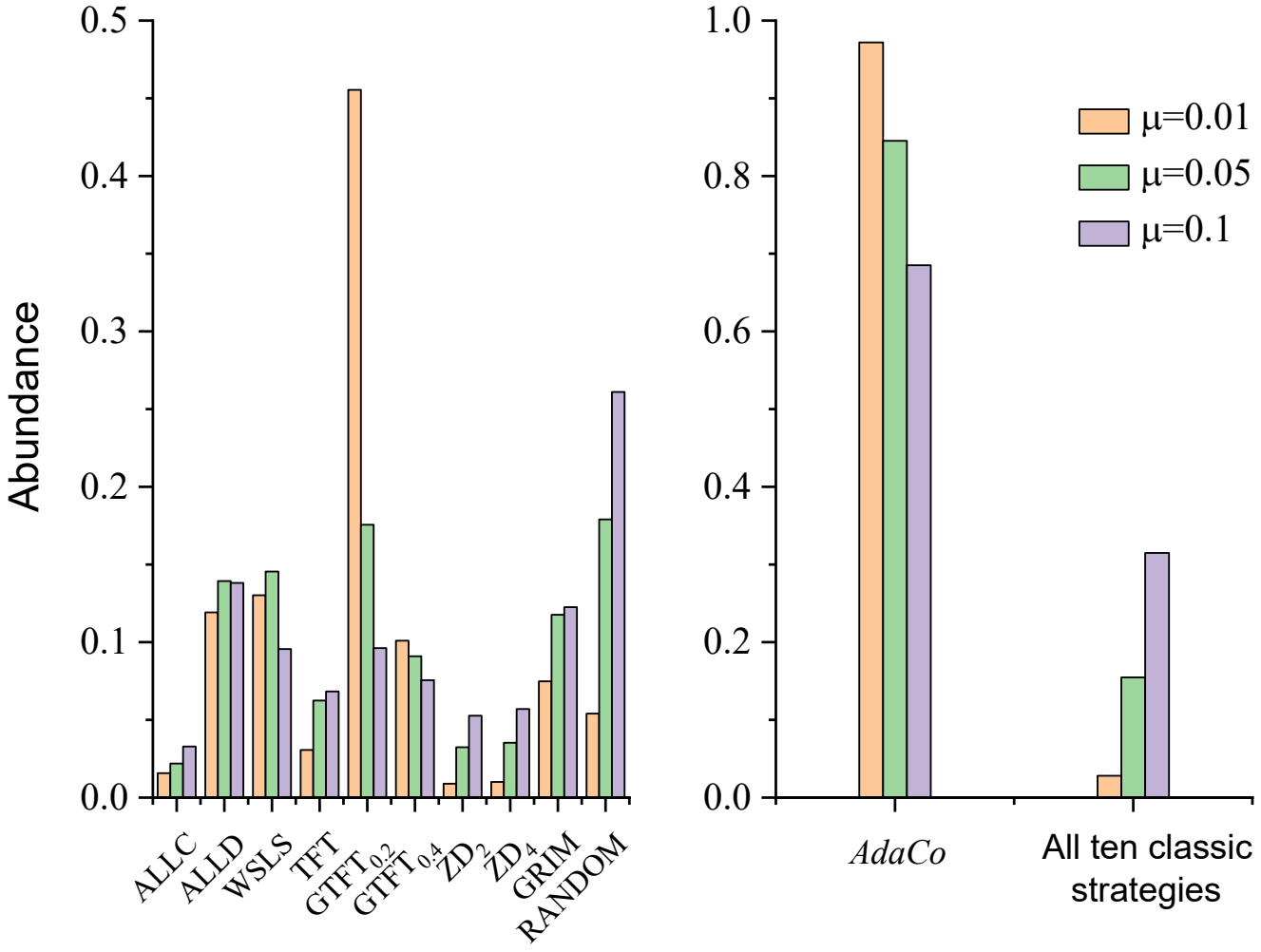


**Figure 2** Evolutionary competition of $AdaCo$ with ten classical strategies under diverse mutation rates. (A) Three mutation rates are examined: $\mu = 0.01$, $\mu = 0.05$, $\mu = 0.1$. we first let these ten classic strategies compete in the evolutionary race. Through computer simulations, we observe that the most prevalent strategy in the population varies with mutation rates, with the $GTFT_{0.2}$ strategy dominating at relatively lower mutation rates ($\mu = 0.01$), while $RANDOM$ dominates at higher mutation rate $\mu = 0.05$ and $\mu = 0.1$. (B) Following the introduction of $AdaCo(K = 3, t = 2)$, it almost entirely dominates the entire population, regardless of whether the mutation rate is high or low. Other parameters include a population size of $M = 100$, $\varepsilon = 0.01$, and selection strength of $\beta = 1$; each simulation is executed for $5 \times 10^8$ mutant strategies.

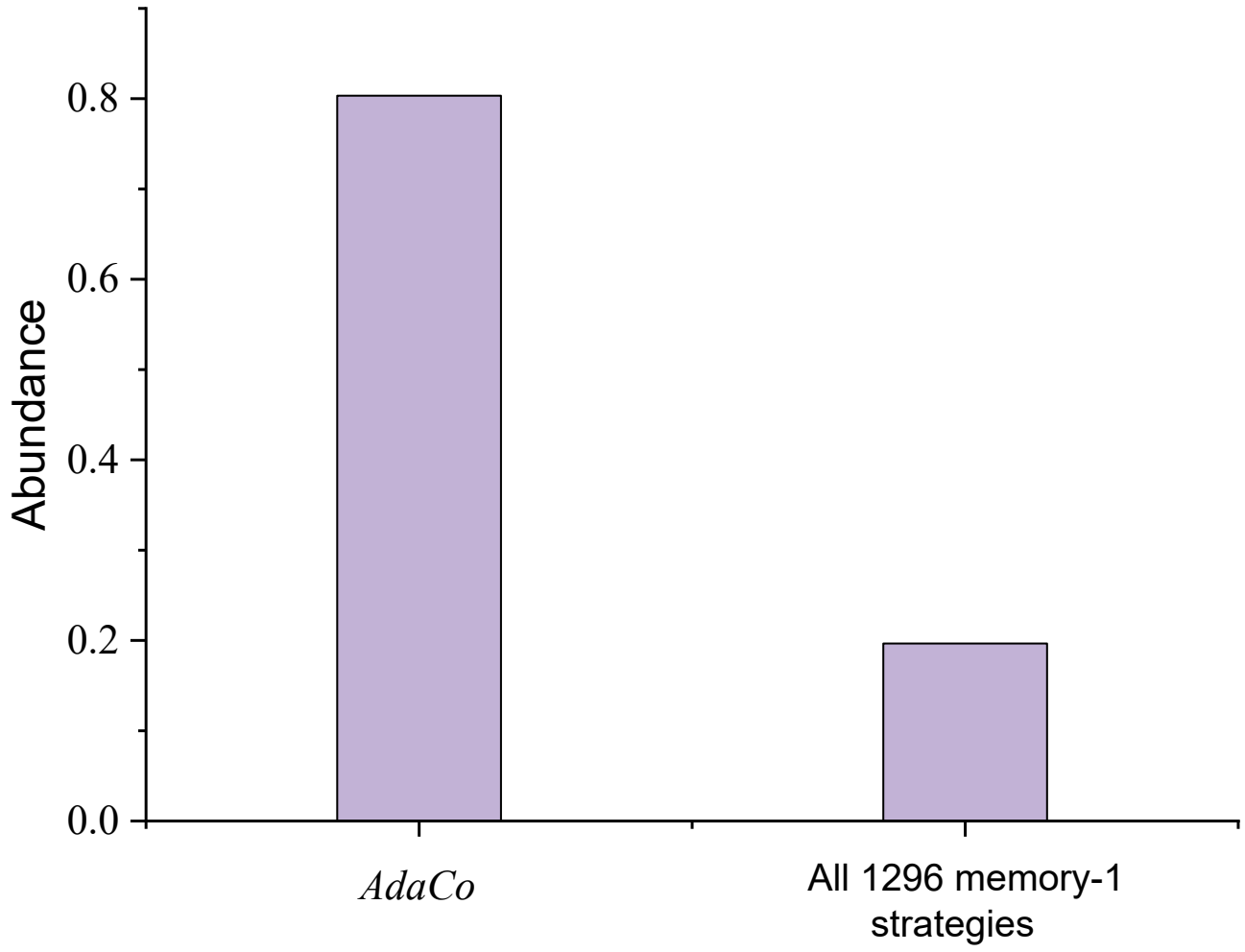


**Figure 3** Evolution of *AdaCo* in the population of memory-1 players under high mutation rates. We considered 1296 memory-1 strategies as described in the main text. Through simulations, we observed that, even at high mutation rates ($\mu = 0.1$), $AdaCo(K = 3, t = 2)$ still becomes predominant, maintaining an average proportion close to 80% throughout the entire evolutionary process time steps in the population. Other parameters include a population size of $M = 100$, $\varepsilon = 0.01$, and selection strength of $\beta = 1$; each simulation is executed for $10^9$ mutant strategies.