-Supplementary Material-

# Evolution of memory strategies in alternating stochastic games


**Jing Zhang[1], Zhihai Rong[1*], Zhi-Xi Wu[2], Bin Xu[3], and Xiaofan Wang[4,5†]**

[1]School of Information and Intelligent Science, Donghua University, Shanghai 201620, China
[2]Institute of Computational Physics and Complex Systems, Lanzhou University, Lanzhou 730000, China
[3]School of Economics, Zhejiang Gongshang University, Hangzhou 310018, China
[4]School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 201620, China
[5]Shanghai Institute of Technology, Shanghai 200235, China
[*]e-mail: rongzhh@gmail.com
[†]e-mail: xfwang@shu.edu.cn


**This PDF file includes:**

- **Supplementary Note 1.** Payoff calculations between pure memory-one strategies in alternating stochastic games.
- **Supplementary Note 2.** Analytical verification of strategy equilibria.
- **Supplementary Note 3.** Impact of decision sequence on evolutionary dynamics in stochastic games.
- **Supplementary Figure S1.** Payoff relationships against random memory-one strategies in alternating stochastic games.
- **Supplementary Figure S2.** Comparison of cooperation rates under alternating and simultaneous interactions.
- **Supplementary Figure S3.** Illustration of simultaneous interactions in stochastic games.
- **Supplementary Figure S4.** Comparison of strategy abundances between alternating and simultaneous stochastic games.
- **Supplementary Table S1–S2.** Payoff matrices of alternating stochastic game with errors.

## Supplementary Note 1. Payoff calculations between pure memory-one strategies in alternating stochastic games

In this section, we introduce the method for calculating the payoff relationships among pure memory-one strategies in repeated alternating games. Let us first consider a single-state game with a $4\times 4$ meta-Markov matrix $\boldsymbol{Q}$, which tracks transitions among the four joint action outcomes $\mathbb{A}=\{CC,CD,DC,DD\}$:

$$\mathbf{Q}=\begin{bmatrix} p_{CC}q_{CC} & p_{CC}(1-q_{CC}) & (1-p_{CC})q_{CD} & (1-p_{CC})(1-q_{CD}) \\ p_{CD}q_{DC} & p_{CD}(1-q_{DC}) & (1-p_{CD})q_{DD} & (1-p_{CD})(1-q_{DD}) \\ p_{DC}q_{CC} & p_{DD}(1-q_{CC}) & (1-p_{DC})q_{CD} & (1-p_{DC})(1-q_{CD}) \\ p_{DD}q_{DC} & p_{DD}(1-q_{DC}) & (1-p_{DD})q_{DD} & (1-p_{DD})(1-q_{DD}) \end{bmatrix}$$

When there is a small implementation error $\varepsilon\to 0$ between strategies, the meta-Markov matrix $\boldsymbol{Q}$ is regular and there exists a unique stationary distribution [1, 2]. Hence, the perturbed matrix $\boldsymbol{Q}_\varepsilon$ can be expanded as:

$$\boldsymbol{Q}_\varepsilon=\boldsymbol{Q}^{(0)}+\varepsilon\boldsymbol{Q}^{(1)}, \tag{S1}$$

where $\boldsymbol{Q}^{(0)}$ is the unperturbed transition matrix and $\boldsymbol{Q}^{(1)}$ is the first-order correction matrix. Similarly, the unique stationary distribution $\boldsymbol{u}_\varepsilon$ is expressed as:

$$\boldsymbol{u}_\varepsilon=\boldsymbol{u}^{(0)}+\varepsilon\boldsymbol{u}^{(1)}, \tag{S2}$$

where $\boldsymbol{u}^{(0)}$ is a stationary distribution of the unperturbed matrix satisfying $\boldsymbol{u}^{(0)}\boldsymbol{Q}^{(0)}=\boldsymbol{u}^{(0)}$, and $\boldsymbol{u}^{(1)}$ is the first-order correction vector. Substituting Eqs. (S1) and (S2) into the stationary condition $\boldsymbol{u}_\varepsilon\boldsymbol{Q}_\varepsilon=\boldsymbol{u}_\varepsilon$ and neglecting the higher-order terms $\mathscr{O}(\varepsilon^2)$, we obtain the perturbation equation as follows:

$$\boldsymbol{u}^{(0)}\boldsymbol{Q}^{(1)}+\boldsymbol{u}^{(1)}\boldsymbol{Q}^{(0)}=\boldsymbol{u}^{(1)}. \tag{S3}$$

As $\varepsilon\to 0$, the unique stationary distribution of the perturbed system converges to a convex combination of the invariant distributions of the unperturbed absorbing classes:

$$\boldsymbol{u}_{\varepsilon\to 0}=\sum_j \alpha_j\boldsymbol{u}_j, \tag{S4}$$

where $\boldsymbol{u}_j$ are the stationary distributions of the absorbing classes of $\boldsymbol{Q}^{(0)}$, and $\alpha_j$ are their corresponding weights derived by solving Eq. (S3). Crucially, the weights $\alpha_j$ derived for the single-state repeated games are also preserved in the two-state stochastic games. Hence, the global stationary distribution $\boldsymbol{v}_{\varepsilon\to 0}$ of the entire $8\times 8$ matrix $\boldsymbol{M}$ is obtained by constructing a linear combination of the invariant distributions of the unperturbed absorbing classes in the full state space, using the same weights $\alpha_j$.

For example, when both players adopt the *TFT* strategy, the perturbation analysis on the $4\times 4$ reduced space yields $\alpha_1=1/2$ for the absorbing class of mutual cooperation $\boldsymbol{u}_1=(1,0,0,0)$ and $\alpha_2=1/2$ for that of mutual defection $\boldsymbol{u}_2=(0,0,0,1)$. To simplify the notation, we denote $\boldsymbol{z}=(z_1,z_2,z_3,z_4)$ corresponds to $\boldsymbol{z}=(z_{CC},z_{CD},z_{DC},z_{DD})$. By assigning these weights to the corresponding unperturbed invariant distributions in the eight-dimensional space of the two-state stochastic game, $\boldsymbol{v}_1=(z_1,0,0,0,1-z_1,0,0,0)$ and $\boldsymbol{v}_2=(0,0,0,z_4,0,0,0,1-z_4)$, we obtain the unique global stationary distribution as:

$$\boldsymbol{v}_{\varepsilon\to 0}=\alpha_1\boldsymbol{v}_1+\alpha_2\boldsymbol{v}_2=(\frac{z_1}{2},0,0,\frac{z_4}{2},\frac{1-z_1}{2},0,0,\frac{1-z_4}{2}) \tag{S5}$$

To accurately calculate the expected payoffs consistent with the random role assignment described in the main text, we evaluate the focal player's outcomes in both the first-mover and second-mover positions. When the focal *TFT* acts as the first player ($X$), its payoff is computed according to Eq. (5) in the main text:

$$\pi^X(\boldsymbol{p}^{TFT},\boldsymbol{q}^{TFT}) = \boldsymbol{v}_{\varepsilon\to 0}\cdot \boldsymbol{E}^X = \frac{z_1\Delta b + b_2 - c}{2}. \tag{S6}$$

Similarly, when acting as the second player ($Y$), its payoff is giving by:

$$\pi^Y(\boldsymbol{q}^{TFT},\boldsymbol{p}^{TFT}) = \boldsymbol{v}_{\varepsilon\to 0}\cdot \boldsymbol{E}^Y = \frac{z_1\Delta b + b_2 - c}{2}. \tag{S7}$$

Therefore, the effective expected payoff is then determined by their weighted average according to Eq. (5) in the main text:

$$\pi(\boldsymbol{p}^{TFT},\boldsymbol{q}^{TFT}) = \frac{1}{2}\pi^X + \frac{1}{2}\pi^Y = \frac{z_1\Delta b + b_2 - c}{2}. \tag{S8}$$

Following this averaging framework for both players, we calculate the effective expected payoff relationships among the sixteen pure memory-one strategies in alternating stochastic games, as presented in Supplementary Tables S1 and S2.

## Supplementary Note 2. Analytical verification of strategy equilibria

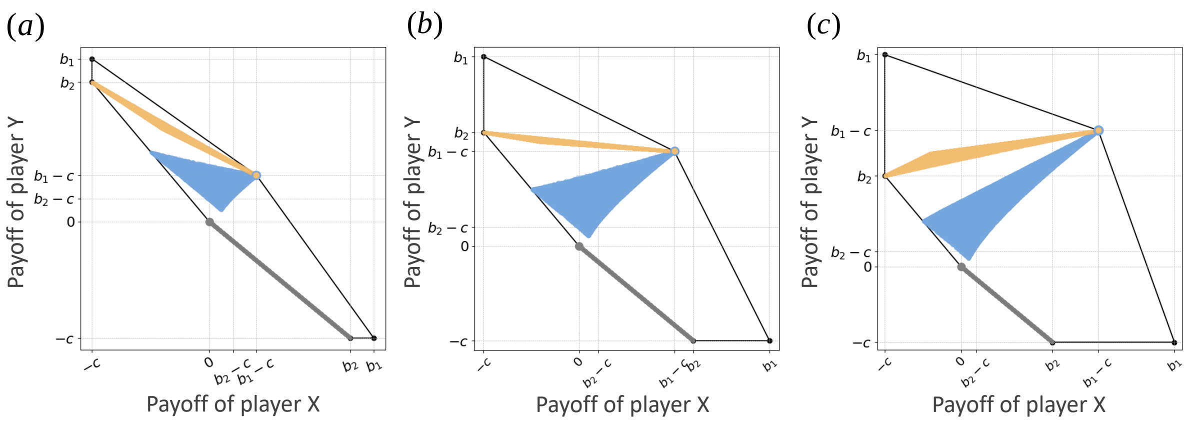


Figure S1: **Payoff relationships against random memory-one strategies in alternating stochastic games.** The environmental difference in (a) $\Delta b = 0.2$, (b) $\Delta b = 0.8$, (c) $\Delta b = 1.6$. The black line bounds the entire feasible region of payoffs. The gray, blue, and orange regions represent the possible payoff pairs for the co-player $Y$ employing any memory-one strategy while player $X$ adopts *AllD*, *FBF* and *AllC*, respectively. The corresponding colored dots indicate the payoffs when both players adopt the same strategy. Other parameters: $b_2 = 1.2$, $c = 1$, $\varepsilon = 0.001$, and $\boldsymbol{z} = (1,0,0,0)$.

To verify the theoretical equilibrium conditions established in the main text, Fig. S1 explores the payoff boundaries of the specific strategies when a focal player $X$ adopts a specific strategy against an opponent $Y$ employing random memory-one strategies for three different stages of $\Delta b$ discussed in the main text.

Since no unilateral deviation by player $Y$ can yield a payoff exceeding that of mutual defection, *AllD* is always a Nash equilibrium. As shown in Fig. S1, the highest gray dots denote the mutual defection payoff, while the lower gray areas confirm that any alternative strategy yields a lower payoff against

Table S1: The payoff matrix of alternating stochastic game with errors.

| | $s_0$ | $s_1$ | $s_2$ | $s_3$ | $s_4$ | $s_5$ | $s_6$ | $s_7$ | $s_8$ |
|---|---|---|---|---|---|---|---|---|---|
| $s_0$ | $\mathbf{0}$ | $\frac{z_4\Delta b+b_2}{2}$ | $0$ | $\frac{z_4\Delta b+b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{6}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $0$ |
| $s_1$ | $-\frac{c}{2}$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{6}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2-2c}{6}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $-\frac{c}{2}$ |
| $s_2$ | $0$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{4}$ | $\mathbf{0}$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{4}$ | $0$ | $\frac{z_2\Delta b+b_2-c}{2}$ | $0$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{4}$ | $0$ |
| $s_3$ | $-\frac{c}{2}$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_2+z_3+2z_4)\Delta b+4b_2-4c}{8}$ | $-\frac{c}{2}$ | $\frac{z_2\Delta b+b_2-c}{2}$ | $\frac{(z_1+z_2+z_3+z_4)\Delta b+4b_2-4c}{8}$ | $\frac{z_4\Delta b+b_2-c}{2}$ | $-\frac{c}{2}$ |
| $s_4$ | $-\frac{c}{3}$ | $\frac{z_4\Delta b+b_2-c}{3}$ | $0$ | $\frac{z_4\Delta b+b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2-2c}{6}$ | $\frac{2z_3\Delta b+2b_2-c}{3}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{3}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $-\frac{c}{3}$ |
| $s_5$ | $-c$ | $-c$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{6}$ | $\frac{(z_2+z_3)\Delta b+2b_2-2c}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $-c$ |
| $s_6$ | $-\frac{c}{2}$ | $-c$ | $0$ | $\frac{(z_2+z_3+2z_4)\Delta b+4b_2-4c}{8}$ | $-\frac{2c}{3}$ | $-c$ | $\frac{(z_1+2z_2+z_3)\Delta b+4b_2-4c}{8}$ | $\frac{(z_1+2z_2+z_3)\Delta b+4b_2-4c}{6}$ | $-\frac{c}{3}$ |
| $s_7$ | $-c$ | $-c$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $-c$ | $-c$ | $\frac{(z_1+2z_2+z_3)\Delta b+4b_2-4c}{6}$ | $\frac{(z_1+2z_2+z_3)\Delta b+4b_2-4c}{6}$ | $-c$ |
| $s_8$ | $0$ | $\frac{z_4\Delta b+b_2}{2}$ | $0$ | $\frac{z_4\Delta b+b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{6}$ | $z_3\Delta b+b_2$ | $\frac{(z_2+z_3)\Delta b+2b_2}{6}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\mathbf{0}$ |
| $s_9$ | $-\frac{c}{2}$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{6}$ | $\frac{(z_3+z_4)\Delta b+2b_2-4c}{6}$ | $\frac{(z_2+z_3+2z_4)\Delta b+4b_2-4c}{8}$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{10}$ | $z_3\Delta b+b_2$ | $\frac{(z_2+2z_3+z_4)\Delta b+4b_2-4c}{8}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\frac{z_1\Delta b+b_2-3c}{5}$ |
| $s_{10}$ | $0$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{4}$ | $0$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{4}$ | $0$ | $\frac{z_2\Delta b+b_2-c}{2}$ | $0$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{z_1\Delta b+b_2-c}{3}$ |
| $s_{11}$ | $-\frac{c}{2}$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_3+z_4)\Delta b+2b_2-4c}{6}$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{4}$ | $-\frac{c}{2}$ | $\frac{z_2\Delta b+b_2-c}{2}$ | $\frac{(z_3+z_4)\Delta b+2b_2-4c}{6}$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{z_1\Delta b+b_2-2c}{3}$ |
| $s_{12}$ | $-\frac{c}{2}$ | $\frac{z_4\Delta b+b_2-2c}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{8}$ | $\frac{(z_2+z_3+2z_4)\Delta b+4b_2-4c}{8}$ | $\frac{(z_2+z_3)\Delta b+2b_2-6c}{12}$ | $\frac{z_3\Delta b+b_2-c}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2-2c}{4}$ | $\frac{3(z_2+z_3)\Delta b+6b_2-4c}{8}$ | $\frac{z_1\Delta b+b_2-3c}{6}$ |
| $s_{13}$ | $-c$ | $-c$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{4}$ | $-c$ | $\frac{z_3\Delta b+b_2-2c}{3}$ | $\frac{3(z_2+z_3)\Delta b+6b_2-8c}{10}$ | $\frac{(z_2+z_3)\Delta b+2b_2-2c}{3}$ | $\frac{z_1\Delta b+b_2-3c}{3}$ |
| $s_{14}$ | $-\frac{2c}{3}$ | $-c$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{6}$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{4}$ | $-\frac{2c}{3}$ | $-c$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{5}$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{4}$ | $\frac{z_1\Delta b+b_2-2c}{3}$ |
| $s_{15}$ | $-c$ | $-c$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{4}$ | $-c$ | $-c$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2-4c}{4}$ | $\frac{z_1\Delta b+b_2-3c}{3}$ |

Table S2: The payoff matrix of alternating stochastic game with errors.

| | $s_9$ | $s_{10}$ | $s_{11}$ | $s_{12}$ | $s_{13}$ | $s_{14}$ | $s_{15}$ |
|---|---|---|---|---|---|---|---|
| $s_0$ | $\frac{z_4\Delta b+b_2}{2}$ | $0$ | $\frac{z_4\Delta b+b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{3}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ |
| $s_1$ | $\frac{(z_2+z_4)\Delta b+2b_2-2c}{6}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2-c}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ |
| $s_2$ | $\frac{(z_1+z_3+2z_4)\Delta b+4b_2-2c}{6}$ | $0$ | $\frac{(z_1+z_3+2z_4)\Delta b+4b_2-2c}{6}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{8}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{4}$ | $\frac{(z_1+z_3)\Delta b+2b_2-c}{3}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{4}$ |
| $s_3$ | $\frac{(z_1+z_2+z_3+z_4)\Delta b+4b_2-4c}{8}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-4c}{8}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{4}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{4}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{4}$ |
| $s_4$ | $\frac{2z_4\Delta b+2b_2-c}{5}$ | $0$ | $\frac{z_4\Delta b+b_2}{2}$ | $\frac{3(z_2+z_3)\Delta b+6b_3-2c}{12}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{3}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ |
| $s_5$ | $-c$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_2+z_3)\Delta b+2b_2-2c}{4}$ | $\frac{2(z_2+z_3)\Delta b+4b_2-2c}{6}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ | $\frac{(z_2+z_3)\Delta b+2b_2}{2}$ |
| $s_6$ | $\frac{(z_1+z_3+2z_4)\Delta b+4b_2-4c}{8}$ | $0$ | $\frac{(z_1+z_3+2z_4)\Delta b+4b_2-2c}{6}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-4c}{8}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-3c}{5}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{5}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{4}$ |
| $s_7$ | $-c$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(z_3+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-6c}{8}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-4c}{6}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{4}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{4}$ |
| $s_8$ | $\frac{(z_1+2z_4)\Delta b+3b_2-c}{5}$ | $\frac{z_1\Delta b+b_2-c}{3}$ | $\frac{(z_1+z_4)\Delta b+2b_2-c}{3}$ | $\frac{(z_1+z_2+z_3)\Delta b+3b_2-c}{6}$ | $\frac{(z_1+z_2+z_3)\Delta b+3b_2-c}{3}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{6}$ | $\frac{(z_1+z_2+z_3)\Delta b+3b_2-c}{3}$ |
| $s_9$ | $\frac{(2z_1+z_2+z_4)\Delta b+4b_2-4c}{8}$ | $z_1\Delta b+b_2-c$ | $z_1\Delta b+b_2-c$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-4c}{8}$ | $\frac{(z_1+z_2+z_3)\Delta b+3b_2-c}{3}$ | $\frac{(4z_1+z_2+z_3)\Delta b+6b_2-4c}{6}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{4}$ |
| $s_{10}$ | $z_1\Delta b+b_2-c$ | $\frac{z_1\Delta b+b_2-c}{2}$ | $z_1\Delta b+b_2-c$ | $\frac{z_1\Delta b+b_2-c}{2}$ | $z_1\Delta b+b_2-c$ | $\frac{2z_1\Delta b+2b_2-2c}{3}$ | $z_1\Delta b+b_2-c$ |
| $s_{11}$ | $z_1\Delta b+b_2-c$ | $z_1\Delta b+b_2-c$ | $\boldsymbol{z_1\Delta b+b_2-c}$ | $\frac{2z_1\Delta b+2b_2-3c}{4}$ | $z_1\Delta b+b_2-c$ | $z_1\Delta b+b_2-c$ | $z_1\Delta b+b_2-c$ |
| $s_{12}$ | $\frac{(z_1+z_4)\Delta b+2b_2-2c}{4}$ | $\frac{z_1\Delta b+b_2-c}{2}$ | $\frac{(2z_1+z_4)\Delta b+3b_2-2c}{4}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-4c}{8}$ | $\frac{(2z_1+3z_2)\Delta b+5b_2-3c}{6}$ | $\frac{(3z_1+2z_2)\Delta b+5b_2-3c}{6}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-2c}{4}$ |
| $s_{13}$ | $\frac{z_1\Delta b+b_2-3c}{3}$ | $z_1\Delta b+b_2-c$ | $z_1\Delta b+b_2-c$ | $\frac{(4z_1+z_2+z_3)\Delta b+6b_2-10c}{12}$ | $\frac{(2z_1+z_2+z_3)\Delta b+4b_2-4c}{6}$ | $\frac{(4z_1+z_2+z_3)\Delta b+6b_2-4c}{6}$ | $\frac{(4z_1+z_2+z_3)\Delta b+6b_2-4c}{6}$ |
| $s_{14}$ | $\frac{2z_1\Delta b+2b_2-3c}{3}$ | $\frac{2z_1\Delta b+2b_2-2c}{3}$ | $z_1\Delta b+b_2-c$ | $\frac{3z_1\Delta b+3b_2-5c}{6}$ | $\frac{2z_1\Delta b+2b_2-3c}{3}$ | $\boldsymbol{z_1\Delta b+b_2-c}$ | $z_1\Delta b+b_2-c$ |
| $s_{15}$ | $\frac{z_1\Delta b+b_2-2c}{2}$ | $z_1\Delta b+b_2-c$ | $z_1\Delta b+b_2-c$ | $\frac{z_1\Delta b+b_2-2c}{2}$ | $\frac{2z_1\Delta b+2b_2-3c}{3}$ | $z_1\Delta b+b_2-c$ | $\boldsymbol{z_1\Delta b+b_2-c}$ |

*AllD*.

Conversely, according to the conditions in Eqs. (1) and (2) of the main text, the cooperative strategies *FBF* and *AllC* may become Nash equilibria, depending on the environmental difference $\Delta b$. If $\Delta b$ is too low to satisfy the condition in Eq. (1), the blue area in Fig. S1(a) demonstrates that *FBF* does not constitute a Nash equilibrium. Specifically, an opponent $Y$ adopting other strategies like *AllD* achieves a higher payoff than the mutual *FBF*, leaving the resident *FBF* population vulnerable to invasion. As $\Delta b$ increases to satisfy Eq. (1), Figs. S1(b) and (c) reveal that mutual *FBF* interactions yield payoffs greater than or equal to those of any deviating strategy, which confirm that *FBF* can resist invasion by defective strategies and maintain cooperation in the population.

Furthermore, the orange dot $(b_1 - c, b_1 - c)$ in Fig. S1 represents the payoff of mutual *AllC* in the good state. Nevertheless, the orange area above this dot in Figs. S1 (a) and (b) indicates that when $\Delta b$ is small and Eq. (2) is not met, *AllC* is susceptible to exploitation by defectors. Once the condition in Eq. (2) is satisfied, Fig. S1(c) confirms that no unilateral deviation by an opponent can improve their payoff against *AllC*, thereby sealing the population against invasion by defective strategies.

## Supplementary Note 3. Impact of decision sequence on evolutionary dynamics in stochastic games.

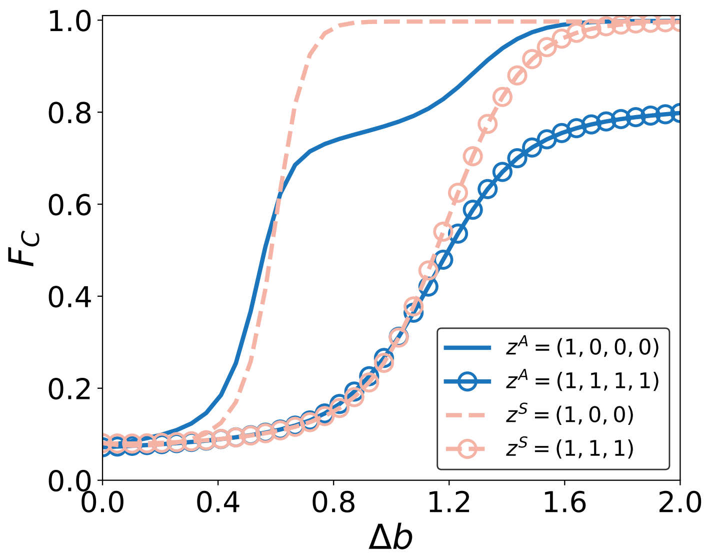


Figure S2: **Comparison of cooperation rates under alternating and simultaneous interactions with rare mutations.** The average population cooperation rate $F_C$ as a function of the environmental difference $\Delta b$, where we compare two-state stochastic games with traditional single-state repeated games under both alternating ($A$) and simultaneous ($S$) mechanisms. Specifically, $\boldsymbol{z}^A = (1,0,0,0)$ and $\boldsymbol{z}^S = (1,0,0)$ denote the alternating and simultaneous stochastic games, respectively. Similarly, $\boldsymbol{z}^A = (1,1,1,1)$ and $\boldsymbol{z}^S = (1,1,1)$ represent the corresponding single-state repeated games in good environment of $b_1 = b_2 + \Delta b$. Parameters: $b_2 = 1.2$, $c = 1$, $\beta = 1$, $N = 100$.

Previous studies have explored how decision-making sequences impact cooperation [3, 4]. It is shown from Fig. S2 that, in traditional single-state repeated games, simultaneous interactions promote cooperation more effectively than alternating ones. This is because simultaneous decisions allow for the emergence of an evolutionarily stable strategy (ESS) under the help of *WSLS* when $b > 2c$ [4]. However, in single-state alternating repeated games, no strategy is strictly stable. When *FBF* is the resident, it is neutrally stable against *ACUD* and *AllC*. Furthermore, these two strategies open an evolutionary backdoor for defective strategies like *AllD* and *Grim*, thereby suppressing the population cooperation rate. To determine whether this established paradigm holds when environmental feedback is introduced, this section investigates the evolutionary dynamics of simultaneous versus alternating interactions in two-state stochastic games.

Since players take actions simultaneously based on the same information, the environmental transition no longer distinguishes the sequence of actions (i.e., *CD* and *DC* are symmetric). Therefore, we adopt a three-dimensional transition rule $\boldsymbol{z}^S = (z_2, z_1, z_0)$, where $z_{n_C}$ represents the probability of transitioning to the good environment given $n_C \in \{0, 1, 2\}$ cooperators in the current round. The sequence in decision-making alters how players utilize historical information, which in turn reshapes their actions and environmental state. As illustrated in Fig. S3, consider the interaction between a player $X$ adopting *TFT* strategy and the co-player $Y$ employing *WSLS* under the transition rule $\boldsymbol{z}^S = (1, 0, 0)$. Initially, both players begin in the good state. If $X$ chooses $C$ and $Y$ selects $D$ in the first round, they will switch to the bad environment ($\boldsymbol{B}$) in the second round. Then $X$ with *TFT* will defect based on $Y$'s previous $D$, while $Y$ with *WSLS* will also defect due to the previous *CD* outcomes. This simultaneous reaction traps them into a cycle among *CD*, *DD*, and *DC* in the bad environmental state, leading to a low long-term payoff, which is different from alternating interactions (Fig. 2(b) in the main text).

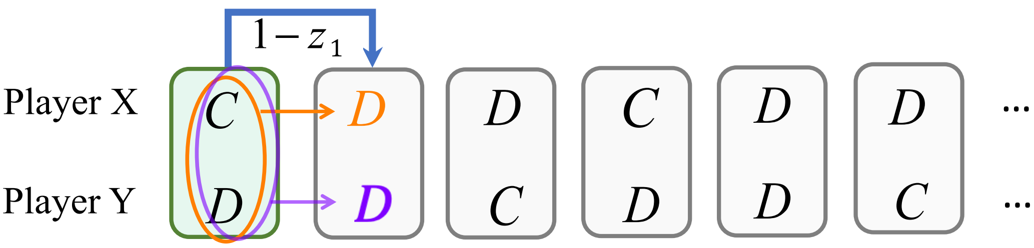


Figure S3: **Illustration of simultaneous interactions in stochastic games.** The environmental state transitions are governed by the rule $\boldsymbol{z}^S = (1, 0, 0)$. A player $X$ adopts the *TFT* strategy, while co-player $Y$ employs *WSLS*. Initially, both players begin in the good state and randomly choose to either cooperate or defect.

Assuming that two players adopt strategies $\boldsymbol{p}$ and $\boldsymbol{q}$, the interactions between the environmental state and the actions of players form a Markov chain $\boldsymbol{M}^S$. We also define $\bar{\boldsymbol{p}} = 1 - \boldsymbol{p}$, $\bar{\boldsymbol{q}} = 1 - \boldsymbol{q}$, and $\bar{\boldsymbol{z}} = 1 - \boldsymbol{z}$. The Markov matrix $\boldsymbol{M}_S$ for simultaneous stochastic games is given by:

$$\mathbf{M}^S = \begin{bmatrix}
p_{CC}q_{CC}z_2 & p_{CC}\bar{q}_{CC}z_2 & \bar{p}_{CC}q_{CC}z_2 & \bar{p}_{CC}\bar{q}_{CC}z_2 & p_{CC}q_{CC}\bar{z}_2 & p_{CC}\bar{q}_{CC}\bar{z}_2 & \bar{p}_{CC}q_{CC}\bar{z}_2 & \bar{p}_{CC}\bar{q}_{CC}\bar{z}_2 \\
p_{CD}q_{DC}z_1 & p_{CD}\bar{q}_{DC}z_1 & \bar{p}_{CD}q_{DC}z_1 & \bar{p}_{CD}\bar{q}_{DC}z_1 & p_{CD}q_{DC}\bar{z}_1 & p_{CD}\bar{q}_{DC}\bar{z}_1 & \bar{p}_{CD}q_{DC}\bar{z}_1 & \bar{p}_{CD}\bar{q}_{DC}\bar{z}_1 \\
p_{DC}q_{CD}z_1 & p_{DC}\bar{q}_{CD}z_1 & \bar{p}_{DC}q_{CD}z_1 & \bar{p}_{DC}\bar{q}_{CD}z_1 & p_{DC}q_{CD}\bar{z}_1 & p_{DC}\bar{q}_{CD}\bar{z}_1 & \bar{p}_{DC}q_{CD}\bar{z}_1 & \bar{p}_{DC}\bar{q}_{CD}\bar{z}_1 \\
p_{DD}q_{DD}z_0 & p_{DD}\bar{q}_{DD}z_0 & \bar{p}_{DD}q_{DD}z_0 & \bar{p}_{DD}\bar{q}_{DD}z_0 & p_{DD}q_{DD}\bar{z}_0 & p_{DD}\bar{q}_{DD}\bar{z}_0 & \bar{p}_{DD}q_{DD}\bar{z}_0 & \bar{p}_{DD}\bar{q}_{DD}\bar{z}_0 \\
p_{CC}q_{CC}z_2 & p_{CC}\bar{q}_{CC}z_2 & \bar{p}_{CC}q_{CC}z_2 & \bar{p}_{CC}\bar{q}_{CC}z_2 & p_{CC}q_{CC}\bar{z}_2 & p_{CC}\bar{q}_{CC}\bar{z}_2 & \bar{p}_{CC}q_{CC}\bar{z}_2 & \bar{p}_{CC}\bar{q}_{CC}\bar{z}_2 \\
p_{CD}q_{DC}z_1 & p_{CD}\bar{q}_{DC}z_1 & \bar{p}_{CD}q_{DC}z_1 & \bar{p}_{CD}\bar{q}_{DC}z_1 & p_{CD}q_{DC}\bar{z}_1 & p_{CD}\bar{q}_{DC}\bar{z}_1 & \bar{p}_{CD}q_{DC}\bar{z}_1 & \bar{p}_{CD}\bar{q}_{DC}\bar{z}_1 \\
p_{DC}q_{CD}z_1 & p_{DC}\bar{q}_{CD}z_1 & \bar{p}_{DC}q_{CD}z_1 & \bar{p}_{DC}\bar{q}_{CD}z_1 & p_{DC}q_{CD}\bar{z}_1 & p_{DC}\bar{q}_{CD}\bar{z}_1 & \bar{p}_{DC}q_{CD}\bar{z}_1 & \bar{p}_{DC}\bar{q}_{CD}\bar{z}_1 \\
p_{DD}q_{DD}z_0 & p_{DD}\bar{q}_{DD}z_0 & \bar{p}_{DD}q_{DD}z_0 & \bar{p}_{DD}\bar{q}_{DD}z_0 & p_{DD}q_{DD}\bar{z}_0 & p_{DD}\bar{q}_{DD}\bar{z}_0 & \bar{p}_{DD}q_{DD}\bar{z}_0 & \bar{p}_{DD}\bar{q}_{DD}\bar{z}_0
\end{bmatrix} \tag{S9}$$

Taking into account a small error rate $\varepsilon$ in the execution of the strategy, we obtain the expected payoffs by calculating the stationary distribution of the matrix $M^S$. Then we focus on the evolutionary dynamics of memory-one strategies both alternating and simultaneous stochastic games.

Figure S2 compares the population cooperation rates $F_C$ across these two interaction mechanisms. Interestingly, introducing stochastic environmental feedback alters the paradigm, yielding a crossover advantage in promoting cooperation that depends on the environmental difference $\Delta b$. When $\Delta b$ is small, alternating stochastic games yield higher cooperation rates than simultaneous ones. To reveal the underlying mechanism, we calculate the stationary abundance distributions of key strategies, as shown in Fig. S4(a). Although *AllD* persists in alternating interactions, the total abundance of defective strategies (*AllD* and *Grim*) is markedly suppressed compared to the simultaneous case. Conversely, alternating interactions enrich the abundance of *TFT* and *FBF*. This advantage stems from the underlying dominance relationships. In alternating stochastic games, when *AllD* is the resident strategy, on the one hand, it is rapidly invaded by *TFT* since the latter dominates former according to the main text. On the other hand, it will be invaded by *eTFT* and *Grim* via neutral drift. More importantly, *TFT* strictly dominates *Grim* provided $z_{CC}\Delta b + b_2 - c > 0$ according to the Tables S1 and S2. This strict dominance provides a greater transition probability from *Grim* to *TFT*, thus increasing the abundance of *TFT*. Consequently, *TFT* acts as a catalyst and enhances a high transition probability to *FBF*, allowing a high abundance of *FBF* although it can be invaded by defective strategies, and then the population has a certain cooperation rate. In contrast, under simultaneous stochastic games, *TFT* only weakly dominates *AllD* and *Grim*. This weaker advantage reduces the transition probability to *TFT*, making it difficult to reach more generous strategies like *WSLS*. Furthermore, *WSLS* is susceptible to be invaded by *AllD* and *Grim* again. As a result, these defective strategies dominate in the population, trapping the population in a low cooperation rate.

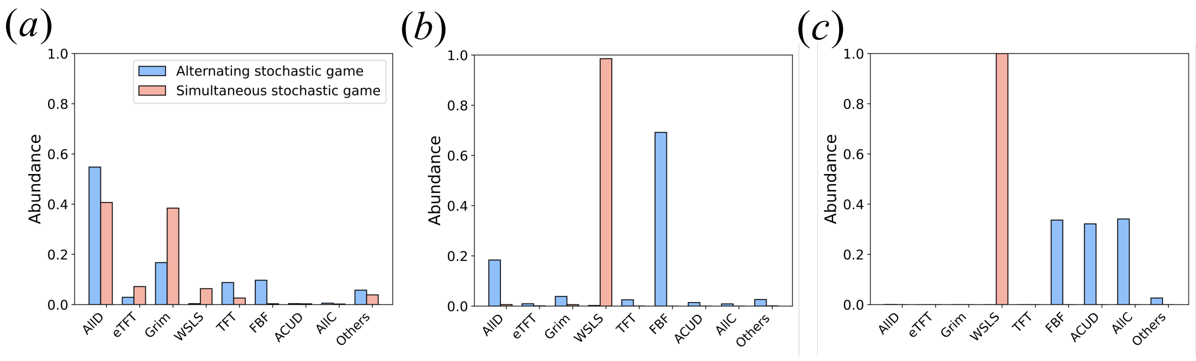


Figure S4: **Comparison of strategy abundances between alternating ($z^A = (1,0,0,0)$) and simultaneous stochastic games ($z^S = (1,0,0)$.** (a) $\Delta b = 0.4$; (b) $\Delta b = 0.8$; (c) $\Delta b = 1.8$. “Others" represents the abundance of the remaining eight memory-one strategies within the strategy space $\mathbb{P}$. Parameters: $b_2 = 1.2$, $c = 1$, $\beta = 1$, $N = 100$.

As $\Delta b$ increases to an intermediate range ($\Delta b \geq 0.4$ when $b_2 = 1.2$ and $c = 1$), both *FBF* and *WSLS* emerge as equilibrium strategies in alternating and simultaneous stochastic games, respectively. However, their evolutionary stability differs. In alternating interactions, the neutral drift among *FBF*, *ACUD* and *AllC* persists as an evolutionary loophole. Since *ACUD* and *AllC* are not yet Nash equilibria at this stage, they remain susceptible to invasion by defective strategies like *AllD* and *Grim*, as shown in Fig. S4(b). In contrast, within simultaneous stochastic games, once *WSLS* becomes the resident strategy under the condition Eq. (1) in the main text, it resists invasions from other strategies [5]. Consequently, simultaneous stochastic games become more effective at stabilizing cooperation than alternating ones Fig. S2.

When $\Delta b$ is sufficiently large ($\Delta b > 1.0$ when $b_2 = 1.2$ and $c = 1$), both alternating and simultaneous interactions in stochastic games achieve almost full cooperation (Fig. S2). In the alternating case, once Eq. (2) in the main text is satisfied, both *ACUD* and *AllC* can resist invasions from defective strategies. Consequently, Fig. S4(c) shows that *FBF*, *ACUD*, and *AllC* form a robust multi-strategy cooperative alliance, driving the population to high cooperation. In contrast, *WSLS* dominates the population and maintains cooperation in simultaneous stochastic games.

In summary, whether alternating or simultaneous interactions is superior in promoting cooperation depends on the environmental feedback and environmental difference. Since alternating and simultaneous mechanisms provide different historical information during the decision-making process of players, the payoff relationship between strategies is reshaped. This sequence interaction ultimately drives the population toward different evolutionary pathways under the distinguishing influence of key cooperative strategies (*FBF* versus *WSLS*).

## Supplementary References


[1] Nowak, M. A. *Evolutionary Dynamics: Exploring the Equations of Life* (Evolutionary Dynamics: Exploring the Equations of Life, 2006).

[2] Sigmund, K. *The calculus of selfishness* (Princeton University Press, 2010).

[3] Nowak, M. A. & Sigmund, K. The alternating prisoner's dilemma. *Journal of Theoretical Biology* **168**, 219–226 (1994).

[4] Park, P. S., Nowak, M. A. & Hilbe, C. Cooperation in alternating interactions with memory constraints. *Nature Communications* **13**, 737 (2022).

[5] Hilbe, C., Šimsa, Š., Chatterjee, K. & Nowak, M. A. Evolution of cooperation in stochastic games. *Nature* **559**, 246–249 (2018).