Supporting Information

Evolutionary dynamics on sequential temporal networks
Anzhi Sheng, Aming Li and Long Wang

1 Evolutionary dynamics

In the main text, we consider two kinds of evolutionary processes on sequential temporal networks. We begin with the first one that the evolution on each snapshot is sufficient. In this case, the fixation of cooperation on a sequential temporal network can be seen as the fixation on each snapshot. Therefore, we separately analyze each snapshot to predict the dynamics on the sequential temporal network.

We consider a population of size $N$. The set of individuals is denoted as $\mathcal{V} = \{1, ..., N\}$. The state of the population is denoted as $x = (x_1, ..., x_N)^T \in \{0, 1\}^N$, where $x_i = 1$ or $x_i = 0$ indicates that the strategy of individual $i$ is cooperation (C) or defection (D). Since the evolutionary process is mutation-free, there are two absorbing states, that is, the population is occupied by cooperators (denoted as $\mathbf{C} = (1, ..., 1)^T$) or defectors (denoted as $\mathbf{D} = (0, ..., 0)^T$).

We follow the workflow proposed by McAvoy & Allen [1] for analyzing the mutation-free evolutionary dynamics under weak selection ($\delta \ll 1$). Following the notation in the main text, each variable under neutral drift and weak selection is labeled with a superscript $\circ$ and $^\ast$, respectively. At each time step, a set of individuals $R \subset \mathcal{V}$ is chosen to be replaced. The replacement is defined by an offspring-to-parent map, $\alpha : R \rightarrow \mathcal{V}$, and $p_{(R,\alpha)}(x)$ denotes the probability of event $(R,\alpha)$ in state $x$. Another useful mapping is the extended version of $\alpha$, denoted as $\tilde{\alpha} : \mathcal{V} \rightarrow \mathcal{V}$, defined by $\tilde{\alpha}(j) = \alpha(j)$ if $j \in R$ and $\tilde{\alpha}(j) = j$ if $j \not\in R$. We denote by $e_{ij}(x)$ the probability of $i$ transmitting its strategy to $j$ in state $x$, given by

$$e_{ij}(x) := \sum_{(R,\alpha)} p_{(R,\alpha)}(x),$$

and its derivative with respect to $\delta$ at $\delta = 0$ is given by

$$\left. \frac{d}{d\delta} \right|_{\delta=0} e_{ij}(x) := \sum_{I \subset \mathcal{V}} c_{ij}^I \prod_{i \in I} x_i.$$

$D_{ij}$ is defined as the degree of the higher-order nonzero term of $c_{ij}^I$, and $D$ is defined as the maximum of $D_{ij}$ for any $i, j$.

Let $\pi_i$ denote the probability of fixation for cooperation when there is initially only a cooperator in site $i$, which is well-known as the reproductive value (RV) of $i$. These probabilities
are the unique solution to the following system of equations
\[
\sum_{j=1}^{N} c_{ij} \pi_{j} = \sum_{j=1}^{N} c_{ji} \pi_{i},
\]
\[
\sum_{j=1}^{N} \pi_{j} = 1.
\]

Then for a given initial configuration \( \xi = (\xi_1, ..., \xi_N)^T \in \{0, 1\}^N \) (except C and D), the fixation probability for this initialization under neutral drift is given by \( \hat{\xi} = \sum_{i=1}^{N} \pi_{i} \xi_{i} \), and the corresponding fixation probability under weak selection is given by
\[
(\rho^{\xi})^* = \hat{\xi} + \delta \left( \sum_{i=1}^{N} \pi_{i} \sum_{t \in \mathcal{V}} c_{i}^{t} \left( \eta_{(i) \cup I}^{\xi} - \eta_{(j) \cup I}^{\xi} \right) \right) + O(\delta^2) \tag{S1}
\]
where \( \eta^{\xi} \) is the unique solution to the equations
\[
\eta_{i}^{\xi} = \hat{\xi} - \xi_{i} + \sum_{(R, a)} p_{(R, a)}^{\xi} \eta_{a(I)}^{\xi} \quad (1 \leq |I| \leq D + 1); \tag{S2a}
\]
\[
\sum_{i=1}^{N} \pi_{i} \eta_{(i)}^{\xi} = 0. \tag{S2b}
\]

For uniform initialization \( \mu \), the fixation probability under weak selection is given by
\[
(\rho^{\mu})^* = \frac{1}{N} + \delta \left( \sum_{i=1}^{N} \pi_{i} \sum_{t \in \mathcal{V}} c_{i}^{t} \left( \eta_{(i) \cup I}^{\mu} - \eta_{(j) \cup I}^{\mu} \right) \right) + O(\delta^2) \tag{S3}
\]
where \( \eta^{\mu} \) is the unique solution to the equations
\[
\eta_{i}^{\mu} = \begin{cases} 
\frac{1}{N} + \sum_{(R, a)} p_{(R, a)}^{\mu} \eta_{a(I)}^{\mu} & 2 \leq |I| \leq D + 1, \\
0 & |I| = 1. \end{cases} \tag{S4}
\]

Let \( \mathcal{T} = (\mathcal{S}, \mathcal{A}) = \{ \mathcal{G}^{(1)}, ..., \mathcal{G}^{(L)} \} \) denote a sequential temporal network, where \( \mathcal{S} \) is the topology when population growth finishes (i.e., the static counterpart), \( \mathcal{A} = \{ \mathbf{a}^{(1)}, ..., \mathbf{a}^{(L)} \} \) is a set of vectors recording the activity of node in each snapshot, \( L \) is the length of the sequential temporal network, and \( \mathcal{G}^{(i)} \) is the \( i \)th snapshot. We consider the evolution on sequential temporal networks under uniform initialization. In this case, the success for the fixation of cooperation in sequential temporal networks requires the fixation of cooperation in each snapshot. Therefore, the fixation probability of sequential temporal networks is given
by
\[ p_T^\mu = \rho_G^{\mu(1)} \prod_{l=2}^L \rho_G^{\mu(l-1)}, \]  

(S5)

2 Specific update rules

A static network \( S \) is represented by a matrix \((w_{ij})_{i,j=1}^N\), where \( w_{ij} \) is the weight of edge \((i,j)\), satisfying \( w_{ij} = w_{ji} \), and \( w_{ij} > 0 \) if an edge exists between individuals \( i \) and \( j \), \( w_{ij} = 0 \) otherwise. The node strength of node \( i \) is \( w_i = \sum_{j=1}^N w_{ij} \), and the probability of a one-step random walk from \( i \) to \( j \) is given by \( p_{ij}^{(1)} = w_{ij}/w_i \). The probability of \( i \) taking \( n \) steps to \( j \) is denoted as \( p_{ij}^{(n)} \), and \( p_{ij}^{(n+m)} = \sum_{k=1}^N p_{ik}^{(n)} p_{kj}^{(m)} \). We note that \( p_{ij}^{(0)} = 1 \) if \( i = j \), otherwise \( p_{ij}^{(0)} = 0 \). The edge-weighted average payoff of node \( i \), depending on the state \( x \), is \( u_i(x) = -cx_i + b \sum_{j=1}^N p_{ij}^{(1)} x_j \). The fitness of node \( i \) is given by \( F_i(x) = 1 + \delta u_i(x) \).

We consider three update rules:

1. Death-Birth (DB) updating [2]: At each time step, an individual \( i \) is selected with a uniform probability to die. A random neighbor \( j \) is selected with probability proportional to \( w_{ji}F_j \) to compete for the empty site.

2. Pairwise-comparison (PC) updating [3]: At each time step, an individual \( i \) is selected with uniform probability and compares payoffs with a neighbor \( j \) selected with probability proportional to \( p_{ij}^{(1)} \). \( j \) reproduces an offspring to replace \( i \) with probability \( \frac{F_j}{F_i + F_j} \).

3. Imitation (IM) updating [2]: At each time step, an individual \( i \) is selected with uniform probability to evaluate its strategy. A random individual \( j \) including \( i \) and its neighbors reproduces an offspring to replace \( i \) with probability proportional to \( w_{ji}F_j \) (here we set \( w_{ii} = 1 \)).

For these update rules, only one individual is replaced at each time step, so the replacement event \((R,a)\) is simplified as \((\{i\}, i \rightarrow j)\), which means that individual \( j \) reproduces an offspring to replace individual \( i \). We note that \( i \) is allowed to equal \( j \).

2.1 DB updating

Under DB updating, the probability of replacement event \((\{j\}, j \rightarrow i)\) in the state \( x \) is given by
\[ p_{(\{j\}, j \rightarrow i)}(x) = \frac{1}{N} \frac{w_{ji}F_j(x)}{\sum_{l=1}^N w_{lj}F_l(x)}. \]
Under neutral drift,
\[ p_{(i,j \rightarrow j)}(x) = \frac{p_{ji}^{(1)}}{N}. \]  
(S6)

The derivative of \( p_{(i,j \rightarrow j)}(x) \) with respect to \( \delta \) at \( \delta = 0 \) is
\[
\left. \frac{d}{d\delta} \right|_{\delta=0} p_{(i,j \rightarrow j)}(x) = \frac{p_{ji}^{(1)}}{N} \left( u_i(x) - \sum_{l=1}^{N} p_{jl}^{(1)} u_l(x) \right) = \sum_{k=1}^{N} c_{ik} x_k,
\]
where
\[
c_{ik} = \frac{p_{ji}^{(1)}}{N} \left( -c \left( p_{ik}^{(0)} - p_{ik}^{(1)} \right) + b \left( p_{ik}^{(1)} - p_{ik}^{(2)} \right) \right).
\]  
(S7)

In this case, the reproductive value of \( i \) is \( \tau_i = w_i / \sum_{k=1}^{N} w_k \). Furthermore, we have the equation
\[
\tau_i p_{ij}^{(1)} = \tau_j p_{ji}^{(1)}.
\]  
(S8)

Substituting Equation S7 into Equation S1 yields
\[
\left( \rho_S^\xi \right)^* = \hat{\xi} + \delta \left( \sum_{i,j,k=1}^{N} \tau_i p_{ij}^{(1)} p_{jk}^{(0)} \eta_{ik}^{\xi} \right) + O(\delta^2).
\]

To simplify notations, let
\[
\eta_{(i)}^{\xi} = \sum_{i,j=1}^{N} \tau_i p_{ij}^{(1)} \eta_{ij}^{\xi}.
\]

Using Equation S8, we have
\[
\begin{align*}
\sum_{i,j,k=1}^{N} \tau_i p_{ij}^{(1)} p_{jk}^{(0)} \eta_{ik}^{\xi} &= \sum_{i,j,k=1}^{N} \tau_i p_{ik}^{(1)} \eta_{ik}^{\xi} = \eta_{(1)}^{\xi}, \\
\sum_{i,j,k=1}^{N} \tau_i p_{ij}^{(1)} p_{ik}^{(0)} \eta_{jk}^{\xi} &= \sum_{j} \tau_j \eta_{jj}^{\xi} = 0, \\
\sum_{i,j,k=1}^{N} \tau_i p_{ij}^{(1)} p_{ik}^{(1)} \eta_{ik}^{\xi} &= \eta_{(1)}^{\xi}, \\
\sum_{i,j,k=1}^{N} \tau_i p_{ij}^{(1)} p_{ik}^{(1)} \eta_{jk}^{\xi} &= \eta_{(2)}^{\xi}, \\
\sum_{i,j,k=1}^{N} \tau_i p_{ij}^{(1)} p_{ik}^{(2)} \eta_{ik}^{\xi} &= \eta_{(2)}^{\xi}, \\
\sum_{i,j,k=1}^{N} \tau_i p_{ij}^{(1)} p_{ik}^{(2)} \eta_{jk}^{\xi} &= \eta_{(1)}^{\xi}, \\
\sum_{i,j,k=1}^{N} \tau_i p_{ij}^{(1)} p_{ik}^{(3)} \eta_{ik}^{\xi} &= \eta_{(3)}^{\xi}, \\
\sum_{i,j,k=1}^{N} \tau_i p_{ij}^{(1)} p_{ik}^{(3)} \eta_{jk}^{\xi} &= \eta_{(1)}^{\xi},
\end{align*}
\]

Therefore,
\[
\left( \rho_S^\xi \right)^* = \hat{\xi} + \frac{\delta}{N} \left( -c \eta_{(2)}^{\xi} + b \left( \eta_{(3)}^{\xi} - \eta_{(1)}^{\xi} \right) \right) + O(\delta^2).
\]
(S9)
We now turn to Equation S2a. For $i \neq j$,
\[
\eta_{ij}^\xi = \tilde{\xi} - \xi_i \xi_j + \frac{1}{N} \sum_{kl=1}^{N} p_{kl}^{(1)} \eta_{kl}^\xi + \frac{1}{N} \sum_{i=1}^{N} \sum_{l=1}^{N} p_{il}^{(1)} \eta_{il}^\xi + \frac{1}{N} \sum_{l=1}^{N} p_{jl}^{(1)} \eta_{jl}^\xi
\]
\[
= \tilde{\xi} - \xi_i \xi_j + \frac{1}{N} \left( \sum_{k \neq l}^{N} p_{kl}^{(1)} \eta_{kl}^\xi + \sum_{i=1}^{N} p_{il}^{(1)} \eta_{il}^\xi + \sum_{l=1}^{N} p_{jl}^{(1)} \eta_{jl}^\xi \right)
\]
which gives
\[
\eta_{ij}^\xi = \frac{N}{2} \left( \tilde{\xi} - \xi_i \xi_j \right) + \frac{1}{2} \sum_{l=1}^{N} \left( p_{il}^{(1)} \eta_{il}^\xi + p_{jl}^{(1)} \eta_{jl}^\xi \right). \tag{S10}
\]

For $i = j$,
\[
\eta_{ii}^\xi = \tilde{\xi} - \xi_i + \frac{1}{N} \sum_{k \neq i}^{N} p_{ki}^{(1)} \eta_{ki}^\xi
\]
\[
= \tilde{\xi} - \xi_i + \frac{1}{N} \left( \sum_{k \neq i}^{N} p_{ki}^{(1)} \eta_{ki}^\xi + \sum_{i=1}^{N} p_{il}^{(1)} \eta_{il}^\xi + \sum_{l=1}^{N} p_{jl}^{(1)} \eta_{jl}^\xi \right)
\]
which gives
\[
\eta_{ii}^\xi = \frac{N}{2} \left( \tilde{\xi} - \xi_i \right) + \sum_{l=1}^{N} p_{il}^{(1)} \eta_{il}^\xi. \tag{S11}
\]

For uniform initialization, let
\[
\eta^\mu_{(k)} = \sum_{i,j=1}^{N} \pi_{ij}^{(k)} \eta_{ij}^\mu,
\]
we have
\[
(p_{S})^{(1)} = \frac{1}{N} + \delta N \left( -c \eta_{(2)}^\mu + b \left( \eta_{(3)}^\mu - \eta_{(1)}^\mu \right) \right) + O(\delta^2), \tag{S12}
\]
and Equation S4 become
\[
\eta_{ij}^\mu = \frac{1}{2} + \frac{1}{2} \sum_{l=1}^{N} \left( p_{il}^{(1)} \eta_{ij}^\mu + p_{jl}^{(1)} \eta_{il}^\mu \right) \quad (i \neq j); \tag{S13}
\]
\[
\eta_{ii}^\mu = 0.
\]
2.2 PC updating

Under PC updating, the probability of replacement event \( \{j\}, j \rightarrow i \) and the marginal probability \( e_{ij} \) in state \( x \) is

\[
p_{\{j\},j\rightarrow i}(x) = e_{ij}(x) = \begin{cases} 
\frac{1}{N} p_{ji}^{(1)} \frac{F_i(x)}{F_i(x) + F_j(x)} & (i \neq j), \\
\frac{1}{N} \sum_{k=1}^{N} p_{jk}^{(1)} \frac{F_j(x)}{F_k(x) + F_j(x)} & (i = j).
\end{cases}
\]

Under neutral drift,

\[
p^\circ_{\{j\},j\rightarrow i}(x) = \begin{cases} 
\frac{1}{2N} p_{ji}^{(1)} & (i \neq j), \\
\frac{1}{2N} & (i = j).
\end{cases}
\]

The derivative of \( p^\circ_{\{j\},j\rightarrow i}(x) \) with respect to \( \delta \) at \( \delta = 0 \) is

\[
\frac{d}{d\delta} \bigg|_{\delta=0} p^\circ_{\{j\},j\rightarrow i}(x) = \begin{cases} 
\frac{1}{N} p_{ji}^{(1)} \left( \frac{u_i(x) - u_j(x)}{4} \right) & (i \neq j), \\
\frac{1}{N} \sum_{k=1}^{N} p_{jk}^{(1)} \left( \frac{u_j(x) - u_k(x)}{4} \right) & (i = j).
\end{cases}
\]

The marginal effect of \( k \) on \( i \) replacing \( j \) is

\[
c_{ik} = \begin{cases} 
\frac{1}{4N} p_{ji}^{(1)} \left( -c(p_{ik}^{(0)} - p_{jk}^{(0)}) \right) + b \left( p_{ik}^{(1)} - p_{jk}^{(1)} \right) & (i \neq j), \\
\frac{1}{4N} \left( -c(p_{jk}^{(0)} - p_{ik}^{(1)}) \right) + b \left( p_{jk}^{(1)} - p_{jk}^{(2)} \right) & (i = j).
\end{cases}
\]

The reproductive value is again \( \pi_i = w_i / \sum_{k=1}^{N} w_k \).

Substituting Equation S17 into Equation S1 yields

\[
\left( \rho^\xi_S \right)^* = \tilde{\xi} + \delta \left( \sum_{i,j,k=1}^{N} \pi_{ij} p_{ji}^{(1)} \left( -c \left( p_{ik}^{(0)} - p_{ik}^{(0)} \right) \right) + b \left( p_{ik}^{(1)} - p_{ik}^{(1)} \right) \right) \left( \eta_{ik} - \eta_{ij} \right) + O(\delta^2).
\]

As before, we have

\[
\left( \rho^\xi_S \right)^* = \tilde{\xi} + \frac{\delta}{N} \left( -c\eta_{(1)} + b \left( \eta_{(1)} - \eta_{(1)} \right) \right) + O(\delta^2).
\]

We then turn to Equation S2a. We see that for \( i \neq j \)

\[
\eta_{ij} = \frac{N}{2} \left( \tilde{\xi} - \xi_{ij} \right) + \frac{1}{2} \sum_{l=1}^{N} \left( p_{ij}^{(1)} \eta_{il} + p_{ji}^{(1)} \eta_{li} \right).
\]
And for $i = j$,
\[
\eta_i^\xi = N \left( \tilde{\xi}_i - \xi_i \right) + \sum_{l=1}^N p_{il}^{(1)} \eta_l^\xi. 
\]  
(S18)

These recurrences are the same as that under DB updating.

For uniform initialization,
\[
(\rho^\mu) = \frac{1}{N} + \frac{\delta}{N} \left( -c \eta^{(1)}_i + b \left( \eta^{(1)}_i - \eta^{(1)}_j \right) \right) + O(\delta^2). 
\]  
(S19)

and the corresponding recurrences are
\[
\eta^\mu_{ij} = \frac{1}{2} + \frac{1}{2} \sum_{l=1}^N \left( p_{il}^{(1)} \eta^\mu_{lj} + p_{jl}^{(1)} \eta^\mu_{il} \right) (i \neq j); \\
\eta^\mu_{ii} = 0. 
\]  
(S20)

### 2.3 IM updating

Under IM updating, the probability of replacement event ($\{j\}, j \rightarrow i$) in the state $x$ is
\[
p_{\{j\}, j \rightarrow i}(x) = \begin{cases} 
1 & \frac{w_{ji} F_i(x)}{N F_j(x) + \sum_{k=1}^N w_{jk} F_k(x)} (i \neq j), \\
\frac{1}{N} & \frac{F_j(x)}{N F_j(x) + \sum_{k=1}^N w_{jk} F_k(x)} (i = j).
\end{cases}
\]

To simplify notations, we define a new matrix \((\tilde{w})_{ij}\) for all $i \neq j$ and $\tilde{w}_{ii} = 1$ for all $i \in \mathcal{V}$. Let $\tilde{p}_{ij} := \tilde{w}_{ij}/\tilde{w}_{i}$ under neutral drift,
\[
p^{\phi}_{\{j\}, j \rightarrow i}(x) = \frac{\tilde{p}_{ji}^{(1)}}{N}.
\]

The derivative of $p^{\phi}_{\{j\}, j \rightarrow i}(x)$ with respect to $\delta$ at $\delta = 0$ is
\[
\frac{d}{d\delta} \bigg|_{\delta=0} p^{\phi}_{\{j\}, j \rightarrow i}(x) = \frac{\tilde{p}_{ji}^{(1)}}{N} \left( u_i(x) - \sum_{l=1}^N \tilde{p}_{jl}^{(1)} u_l(x) \right). 
\]  
(S21)

The marginal effect of $k$ on $i$ replacing $j$ is
\[
\epsilon_{ik}^{ij} = \frac{\tilde{p}_{ji}^{(1)}}{N} \left( -c \left( p_{ik}^{(0)} - \tilde{p}_{ik}^{(1)} \right) + b \left( p_{ik}^{(1)} - \sum_{m=1}^N \tilde{p}_{jm}^{(1)} p_{mk}^{(1)} \right) \right). 
\]  
(S22)
The reproductive value is \( \pi_i = \bar{w}_i / \sum_{k=1}^{N} \bar{w}_k \). We also have the equation

\[
\pi_i \tilde{p}_{ij}^{(1)} = \pi_j \tilde{p}_{ji}^{(1)} \tag{S23}
\]

Substituting Equation S22 into Equation S1 yields

\[
\left( \rho_S^\xi \right)^* = \hat{\xi} + \delta \left( -c \left( p_{(0)}^{(0)} - \tilde{p}_{(1)}^{(1)} \right) + b \left( p_{(1)}^{(1)} - \sum_{m=1}^{N} \tilde{p}_{(1)}^{(1)} p_{(1)}^{(1)} m k \right) \right) \left( \eta_{(2,0)}^{\xi} - \eta_{(0,1)}^{\xi} \right) + O \left( \delta^2 \right) .
\]

Let

\[
\tilde{\eta}_{(m,n)}^{\xi} = \sum_{i,j,k=1}^{N} \pi_i \tilde{p}_{ik}^{(m)} p_{kj}^{(n)} \eta_{ij}^{\xi} \quad \tilde{\eta}_{(m,n)}^{\mu} = \sum_{i,j,k=1}^{N} \pi_i \tilde{p}_{ik}^{(m)} p_{kj}^{(n)} \eta_{ij}^{\mu}.
\]

We obtain

\[
\left( \rho_S^\xi \right)^* = \hat{\xi} + \delta \left( -c \tilde{\eta}_{(2,0)}^{\xi} + b \left( \tilde{\eta}_{(2,1)}^{\xi} - \tilde{\eta}_{(0,1)}^{\xi} \right) \right) + O \left( \delta^2 \right) . \tag{S24}
\]

Similarly, for \( i \neq j \)

\[
\eta_{ij}^{\xi} = \frac{N}{2} \left( \tilde{\xi} - \xi_{ij} \right) + \frac{1}{2} \sum_{l=1}^{N} \left( \tilde{p}_{il}^{(1)} \eta_{lj}^{\xi} + \tilde{p}_{jl}^{(1)} \eta_{li}^{\xi} \right) . \tag{S25}
\]

And for \( i = j \),

\[
\eta_{ii}^{\xi} = N \left( \tilde{\xi} - \xi_i \right) + \sum_{l=1}^{N} \tilde{p}_{il}^{(1)} \eta_{li}^{\xi} . \tag{S26}
\]

For uniform initialization,

\[
\left( \rho_S^\mu \right)^* = \hat{\xi} + \delta \left( -c \tilde{\eta}_{(2,0)}^{\mu} + b \left( \tilde{\eta}_{(2,1)}^{\mu} - \tilde{\eta}_{(0,1)}^{\mu} \right) \right) + O \left( \delta^2 \right) . \tag{S27}
\]

and the corresponding recurrences are

\[
\eta_{ij}^{\mu} = 1 + \frac{1}{2} \sum_{l=1}^{N} \left( \tilde{p}_{il}^{(1)} \eta_{lj}^{\mu} + \tilde{p}_{jl}^{(1)} \eta_{li}^{\mu} \right) \quad (i \neq j) ;
\]

\[
\eta_{ii}^{\mu} = 0 . \tag{S28}
\]

3 Condition for promotion on sequential temporal networks

In this section, we compare the fixation probability of sequential temporal networks and corresponding static networks under these three updating rules.
3.1 Neutral drift

We notice that under any update rule,

\[
\rho_S^\circ \circ = \hat{\xi} = \frac{\sum_{i=1}^{N} w_i \xi_i}{\sum_{i=1}^{N} w_i}, \quad \rho_S^\circ = \frac{1}{N}.
\]

For a sequential temporal network \( T = (S, A) = \{G^{(1)}, ..., G^{(L)}\} \), the fixation probability under neutral drift is given by

\[
\rho_T^\circ = \frac{1}{\sum_{i=1}^{N} a_i^{(1)} \prod_{l=2}^{L} \sum_{i=1}^{N} \pi_i|G^{(l)}|a_i^{(l-1)}},
\]

where \( \pi_i|G^{(l)} \) represents \( \pi_i \) in \( G^{(l)} \).

3.1.1 DB and PC updating

The reproductive value of \( i \) is the same under DB and PC updating, given by \( \hat{\xi} = \sum_{i=1}^{N} w_i \xi_i / \sum_{i=1}^{N} w_i \). So a node with a higher node strength has a greater reproductive value and therefore greater influence on the evolutionary dynamics.

We first focus on sequential temporal networks with \( L = 2 \) (i.e., \( T = \{G^{(1)}, G^{(2)}\} \) and \( S = G^{(2)} \)). We assume that \( G^{(1)} \) \( (G^{(2)} \) has \( m \) \((m + \Delta m) \) nodes and the average degree of \( G^{(1)} \) \( (G^{(2)} \) is \( k_1 \) \((k_2 \). Furthermore, we assume that there are no edges among the newly added nodes \( \Delta m \) in \( G^{(2)} \).

The equivalent condition of \( \rho_T^\circ \circ > \rho_S^\circ \circ \) is

\[
\frac{1}{m} \frac{mk_1 + (m + \Delta m)k_2}{2(m + \Delta m)k_2} > \frac{1}{m + \Delta m}.
\]

Equation S29 is equivalent to

\[
m(m + \Delta m)k_1 - k_2(m + \Delta m)(m - \Delta m) > 0.
\]

It is straightforward to obtain the first condition in the main text (Equation 6a),

\[
\Delta m \geq m.
\]

For the second condition, we notice that

\[
mk_1 + 2\Delta K = k_2(m + \Delta m),
\]

where \( \Delta K \) is the number of newly added edges. Combined with Equation S30, when \( \Delta m < m \),
we obtain the second condition
\[ \Delta K < \frac{1}{2} \left( \frac{m(m + \Delta m)k_1}{m - \Delta m} - mk_1 \right) = \frac{m\Delta mk_1}{m - \Delta m}. \] (S32)
When \( m \ll \Delta m \), the condition is transformed into
\[ \Delta K < \Delta mk_1. \]

Moreover, we study three variations of the model: the rule of increasing the fixation probability for static networks, the rule in weighted networks, and the rule when internal edges exist in newly added nodes.

**Variation 1:** From another perspective, for a given static network \( S \), we can make the fixation probability larger by constructing a sequential temporal network. We assume that \( G(1) (G(2)) \) has \( m - \Delta m \) nodes and the average degree of \( G(1) (G(2)) \) is \( k_1 (k_2) \). There are no edges between the deleted nodes \( \Delta m \) in \( S \). Similarly, we get the equivalent condition for promotion
\[ \frac{1}{m - \Delta m} \frac{(m - \Delta m)k_1 + mk_2}{2mk_2} > \frac{1}{m}. \] (S33)
We denote by \( \Delta K \) the number of edges removed from \( S \), then Equation S33 is equivalent to
\[ \Delta K < \Delta mk_2, \]
meaning that the number of deleted edges should be less than the number of edges \((\Delta mk_2)\) formed by the number of deleted nodes under the average degree of \( S \).

**Variation 2:** We assume that the average weight of \( G(1) (G(2)) \) is \( w_1 (w_2) \). The conditions for cooperation are
\( (i) \) \( \Delta m \geq m \),
\( (ii) \) \( \Delta m < m \), \( \Delta K < \frac{m\Delta mw_1}{m - \Delta m} \).

**Variation 3:** We assume that there are \( g_1 \) internal edges nodes and \( g_2 \) external edges connecting to \( G(1) \) in \( \Delta m \) newly added nodes. The upper bound of \( g_1 \) is \( \Delta m(\Delta m - 1)/2 \).

Similar to Equations S31 and S32, the condition for promotion is
\[ \frac{1}{m \cdot mk_1 + g_2} \frac{mk_1 + g_2}{2g_1 + g_2} > \frac{1}{m + \Delta m} \Leftrightarrow m\Delta mk_1 + (\Delta m - m)g_2 - 2mg_1 > 0. \] (S34)
When \( g_1 > \Delta mk_1/2 \), meaning that new nodes are highly connected with each other, the condition becomes
\[ \Delta m > m, \quad g_2 > \frac{2mg_1 - m\Delta mk_1}{\Delta m - m}. \]
When \( g_1 < \Delta m k_1 / 2 \), the condition becomes

\[ \Delta m \geq m, \]

or

\[ \Delta m < m, \quad g_2 < \frac{m \Delta m k_1 - 2 m g_1}{m - \Delta m}. \]

Especially, when \( g_1 = 0 \), the condition degenerates to Equations S33 and S34.

When newly joined nodes have no internal edges with each other in each snapshot, the necessary and sufficient condition of \( (\rho_T^\mu) > (\rho_S^\mu)^\circ \) for any \( L \geq 3 \) is

\[
\frac{1}{m_1} \prod_{l=2}^{L} \frac{m_{l-1} k_{l-1} + m_l k_l}{2m_l k_l} > \frac{1}{m_1 k_L},
\]

where \( m_l \) and \( k_l \) are the number of nodes and average connectivity of \( G^{(l)} \), respectively. In particular, using mathematical induction, a sufficient condition of \( (\rho_T^\mu) > (\rho_S^\mu)^\circ \) for any \( L \geq 3 \) is every pair of adjacent snapshots satisfies Equation S33 or Equation S34.

3.1.2 IM updating

Here, we turn to the IM updating. The fixation probability \( (\rho_S^\xi)^\circ \) is given by

\[
\hat{\xi} = \frac{\sum_{i=1}^{N} \tilde{w}_i \xi_i}{\sum_{i=1}^{N} \tilde{w}_i} = \frac{\sum_{i=1}^{N} (w_i + 1) \xi_i}{\sum_{i=1}^{N} w_i + N}.
\]

We begin with \( L = 2 \). The notations of sequential temporal networks are the same as that in section 3.1.1. The equivalent condition of \( (\rho_T^\mu) > (\rho_S^\mu)^\circ \) is

\[
\frac{1}{m} \frac{m(k_1 + 1) + ((m + \Delta m)k_2 - mk_1)/2}{(m + \Delta m)(k_2 + 1)} > \frac{1}{m + \Delta m},
\]

which gives

\[
(m + \Delta m) \Delta K + m \Delta mk_1 > 2m \Delta K.
\]

This condition is the same as that in section 3.1.1.

For \( L \geq 3 \), the necessary and sufficient condition of \( (\rho_T^\mu) > (\rho_S^\mu)^\circ \) is

\[
\frac{1}{m_1} \prod_{l=2}^{L} \frac{m_{l-1} k_{l-1} + m_l k_l + 2m_{l-1}}{2m_l (k_l + 1)} > \frac{1}{m_1 k_L}.
\]
3.2 Weak selection

Using mathematical induction, we can demonstrate that when the relation \((\rho_T^\mu)^\circ > (\rho_S^\mu)^\circ\) holds, the relation \((\rho_T^\mu)^* > (\rho_S^\mu)^*\) also holds. Therefore, we consider the equivalent condition for promotion under weak selection only when \((\rho_T^\mu)^\circ = (\rho_S^\mu)^\circ\). In this case, we compare the first-order term of \((\rho_T^\mu)^*\) and \((\rho_S^\mu)^*\), and ignore the higher-order terms.

We start with \(L = 2\). The first-order coefficient of two types of fixation probabilities on snapshot \(G(i)\) is given by

\[
g(G(i); \xi) = \begin{cases} 
\frac{1}{N} \left( -c \eta(1) + b (\eta(2) - \eta(1)) \right) & \text{(DB updating),} \\
\frac{1}{N} \left( -c \eta(1) + b (\eta(2) - \eta(1)) \right) & \text{(PC updating),} \\
\frac{1}{N} \left( -c \eta(2) + b (\eta(2) - \eta(1)) \right) & \text{(IM updating).}
\end{cases}
\]

\[
g(G(i); \mu) = \begin{cases} 
\frac{1}{N} \left( -c \eta(2) + b (\eta(3) - \eta(1)) \right) & \text{(DB updating),} \\
\frac{1}{N} \left( -c \eta(1) + b (\eta(2) - \eta(1)) \right) & \text{(PC updating),} \\
\frac{1}{N} \left( -c \eta(2) + b (\eta(2) - \eta(1)) \right) & \text{(IM updating).}
\end{cases}
\]

and the condition of \((\rho_T^\mu)^* > (\rho_S^\mu)^*\) is

\[
\frac{1}{N} \Big| G(i) \Big| g(G(2); a(1)) + \hat{\xi} \Big| G(i) \Big| g(G(1); \mu) > g(G(2); \mu),
\]

(S37)

where \(\frac{1}{N} \Big| G(i) \Big|\) means the size of snapshot \(G(i)\), and \(\hat{\xi} \Big| G(i) \Big|\) means the RV-weighted value of \(\hat{\xi}\) on snapshot \(G(i)\).

For \(L \geq 3\), let \(A_1 = (1/N) \Big| G(1) \Big|\) and \(A_i = a^{(i-1)} \Big| G(i) \Big|\) \((i = 2, ..., L)\), the necessary and sufficient condition of \((\rho_T^\mu)^* > (\rho_S^\mu)^*\) is given by

\[
\tilde{A}_1 g(G(1); \mu) + \sum_{i=2}^{L} \tilde{A}_i g(G(i); a^{(i-1)}) > g(G(L); \mu),
\]

(S38)

where \(\tilde{A}_i = \prod_{j \neq i} A_j\). Once again, when each pair of adjacent snapshots of \(T\) satisfies Equation S37, we also have \((\rho_T^\mu)^* > (\rho_S^\mu)^*\).
4 Critical benefit-to-cost ratio

We also use the critical benefit-to-cost ratio \( \left( \frac{b}{c} \right)^* \) to compare the cooperation-promoting effect of sequential temporal networks and static networks. We say a sequential temporal network \( \mathcal{T} \) promotes cooperation if

\[
\begin{align*}
\left( \frac{b}{c} \right)^*_S \mu & > \left( \frac{b}{c} \right)^*_T \mu > 0, \quad \text{(S39a)} \\
\left( \frac{b}{c} \right)^*_S \mu & < 0 < \left( \frac{b}{c} \right)^*_T. \quad \text{(S39b)}
\end{align*}
\]

For an arbitrary initial configuration \( \xi \), the critical benefit-to-cost ratio of \( S \) is given as

\[
\left( \frac{b}{c} \right)^*_S \xi = \begin{cases} 
\frac{\eta^{\xi}_2}{\eta^{\xi}_3 - \eta^{\xi}_1} & \text{(DB updating)}, \\
\frac{\eta^{\xi}_1}{\eta^{\xi}_2 - \eta^{\xi}_1} & \text{(PC updating)}, \\
\frac{\tilde{\eta}^{\xi}_2}{\tilde{\eta}^{\xi}_{(2,1)} - \tilde{\eta}^{\xi}_{(0,1)}} & \text{(IM updating)}.
\end{cases}
\]

For uniform initialization, the critical benefit-to-cost ratio is

\[
\left( \frac{b}{c} \right)^*_S \mu = \begin{cases} 
\frac{\eta^{\mu}_2}{\eta^{\mu}_3 - \eta^{\mu}_1} & \text{(DB updating)}, \\
\frac{\eta^{\mu}_1}{\eta^{\mu}_2 - \eta^{\mu}_1} & \text{(PC updating)}, \\
\frac{\tilde{\eta}^{\mu}_2}{\tilde{\eta}^{\mu}_{(2,1)} - \tilde{\eta}^{\mu}_{(0,1)}} & \text{(IM updating)}.
\end{cases}
\]

For a sequential temporal network $T$ with $L \geq 2$, the critical benefit-to-cost ratio is

$$\left( \begin{array}{c}
\tilde{A}_1 \left( \frac{\eta^u}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2)} - \eta^{(i)}_{(1)}}{N} \right) \right)_{G^{(i)}} \\
\tilde{A}_1 \left( \frac{\eta^{(u)}_{(2)} - \eta^{(u)}_{(1)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2)} - \eta^{(i)}_{(1)}}{N} \right) \right)_{G^{(i)}} \\
\tilde{A}_1 \left( \frac{\eta^u_{(2)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2)}}{N} \right) \right)_{G^{(i)}} \\
\tilde{A}_1 \left( \frac{\eta^u_{(2,1)} - \eta^u_{(0,1)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2,1)} - \eta^{(i)}_{(0,1)}}{N} \right) \right)_{G^{(i)}} \\
\end{array} \right)$$

(DB updating)

$$\left( \begin{array}{c}
\tilde{A}_1 \left( \frac{\eta^u_{(2,1)} - \eta^u_{(0,1)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2,1)} - \eta^{(i)}_{(0,1)}}{N} \right) \right)_{G^{(i)}} \\
\tilde{A}_1 \left( \frac{\eta^{(u)}_{(2,1)} - \eta^{(u)}_{(0,1)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2,1)} - \eta^{(i)}_{(0,1)}}{N} \right) \right)_{G^{(i)}} \\
\tilde{A}_1 \left( \frac{\eta^{(u)}_{(2,0)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2,0)}}{N} \right) \right)_{G^{(i)}} \\
\tilde{A}_1 \left( \frac{\eta^{(u)}_{(2,1)} - \eta^{(u)}_{(1,1)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2,1)} - \eta^{(i)}_{(1,1)}}{N} \right) \right)_{G^{(i)}} \\
\end{array} \right)$$

(PC updating) (S42)

$$\left( \begin{array}{c}
\tilde{A}_1 \left( \frac{\eta^u_{(2,1)} - \eta^u_{(0,1)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2,1)} - \eta^{(i)}_{(0,1)}}{N} \right) \right)_{G^{(i)}} \\
\tilde{A}_1 \left( \frac{\eta^{(u)}_{(2,1)} - \eta^{(u)}_{(0,1)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2,1)} - \eta^{(i)}_{(0,1)}}{N} \right) \right)_{G^{(i)}} \\
\tilde{A}_1 \left( \frac{\eta^{(u)}_{(2,0)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2,0)}}{N} \right) \right)_{G^{(i)}} \\
\tilde{A}_1 \left( \frac{\eta^{(u)}_{(2,1)} - \eta^{(u)}_{(1,1)}}{N} \right) \bigg|_{G^{(i)}} + \left( \sum_{i=2}^{L} \tilde{A}_i \left( \frac{\eta^{(i-1)}_{(2,1)} - \eta^{(i)}_{(1,1)}}{N} \right) \right)_{G^{(i)}} \\
\end{array} \right)$$

(IM updating)

where $\tilde{A}_i$ is the same as that in Equation S38. Substituting Equations S40 and S42 into Equation S39 yields a general condition about the critical ratio.

5 Mean-field approximation

In this section, we develop the mean-field approximation proposed in Ref. [4] to calculate the fixation probability and the critical benefit-to-cost ratio under weak selection.

We begin with DB and PC updating. For any arbitrary initial configuration $\xi = (\xi_1, ..., \xi_N)$ except $C, D$, let $t_{ij} = \frac{N}{2}(\xi_i - \xi_j^c)$ $(i \neq j)$ and $t_{ii} = 1/2$ for all $i$. We have a recurrence relation
for $\eta^{(n)}_{(n)}$: \[ \eta^{(n)}_{(n)} = \sum_{i,j=1}^{N} \pi_i p_{ij}^{(n)} \eta_{ij}^{(n)} \]

$$= \sum_{i,j=1}^{N} \pi_i p_{ij}^{(n)} \left( t_{ij} + \frac{1}{2} \sum_{k=1}^{N} \left( p_{ik} \eta_{kj}^{(n)} + p_{jk} \eta_{ik}^{(n)} \right) \right) + \sum_{i=1}^{N} \pi_i p_{ii}^{(n)} \eta_{ii}^{(n)}$$

$$= \sum_{i,j=1}^{N} \pi_i p_{ij}^{(n)} \left( t_{ij} + \frac{1}{2} \sum_{k=1}^{N} \left( p_{ik} \eta_{kj}^{(n)} + p_{jk} \eta_{ik}^{(n)} \right) \right) + \sum_{i=1}^{N} \pi_i p_{ii}^{(n)} \eta_{ii}^{(n)}$$

$$= \sum_{i=1}^{N} \pi_i p_{ii}^{(n)} \left( t_{ii} + \frac{1}{2} \sum_{k=1}^{N} (p_{ik} \eta_{ki}^{(n)} + p_{ik} \eta_{ik}^{(n)}) \right) \quad \text{for } \eta^{(n)}_{(n)}$$

To further simplify notations, let $q_i = \sum_{k=1}^{N} (p_{ik} \eta_{ik}^{(n)} - \eta_{ii}^{(n)}) + t_{ii}$, the recurrence relation is simplified as

$$\eta^{(n+1)}_{(n)} = \eta^{(n)}_{(n)} + \sum_{i=1}^{N} \pi_i p_{ii}^{(n)} q_i - \sum_{i,j=1}^{N} \pi_i p_{ij}^{(n)} t_{ij}. \quad (S43)$$

In particular, we have

$$\eta^{(0)}_{(0)} = \sum_{i,j=1}^{N} \pi_i p_{ij}^{(0)} \eta_{ij} = \sum_{i=1}^{N} \pi_i \eta_{ii} = 0,$$

$$\eta^{(1)}_{(1)} = \sum_{i=1}^{N} \pi_i p_{ii}^{(0)} q_i - \sum_{i,j=1}^{N} \pi_i p_{ij}^{(0)} t_{ij}, \quad (S44)$$

$$\eta^{(2)}_{(2)} = \sum_{i=1}^{N} \pi_i q_i (1 + p_{ii}^{(1)}) - \sum_{i,j=1}^{N} \pi_i t_{ij} (p_{ij}^{(0)} + p_{ij}^{(1)}),$$

$$\eta^{(3)}_{(3)} = \sum_{i=1}^{N} \pi_i q_i (1 + p_{ii}^{(1)} + p_{ii}^{(2)}) - \sum_{i,j=1}^{N} \pi_i t_{ij} (p_{ij}^{(0)} + p_{ij}^{(1)} + p_{ij}^{(2)}).$$

Similarly, let $\tilde{q}_i = \sum_{k=1}^{N} (\tilde{p}_{ik} \eta_{ik}^{(n)} - \eta_{ii}^{(n)}) + t_{ii}$, the recurrence relation under IM updating is

$$\eta^{(n+1),0}_{(n,0)} = \eta^{(n),0}_{(n,0)} + \sum_{i=1}^{N} \pi_i \tilde{p}_{ii}^{(n)} \tilde{q}_i - \sum_{i,j=1}^{N} \pi_i \tilde{p}_{ij}^{(n)} t_{ij}, \quad (S45)$$

15
\[\eta_\xi(0,0) = 0,\]
\[\eta_\xi(1,0) = \sum_{i=1}^{N} \tau_i \tilde{q}_i - \sum_{i,j=1}^{N} \tau_i \tilde{p}_{ij} t_{ij},\]
\[\eta_\xi(2,0) = \sum_{i=1}^{N} \tau_i \tilde{q}_i (1 + \tilde{p}_{ii}^{(1)}) - \sum_{i,j=1}^{N} \tau_i t_{ij} (\tilde{p}_{ij}^{(0)} + \tilde{p}_{ij}^{(1)}),\]
\[\eta_\xi(3,0) = \sum_{i=1}^{N} \tau_i \tilde{q}_i (1 + \tilde{p}_{ii}^{(1)} + \tilde{p}_{ii}^{(2)}) - \sum_{i,j=1}^{N} \tau_i t_{ij} (\tilde{p}_{ij}^{(0)} + \tilde{p}_{ij}^{(1)} + \tilde{p}_{ij}^{(2)}).\] (S46)

Since \(S\) is connected, random walks on \(S\) are ergodic [5]. Therefore, the equation \(\lim_{n \to \infty} p_n^{(n)} = \tau_i (\lim_{n \to \infty} \tilde{p}_{ij}^{(n)} = \tau_i)\) holds for all \(j \in S\) and \(\lim_{n \to \infty} \eta_\xi(n) = \eta_\xi, (\lim_{n \to \infty} \eta_\xi(n,0) = \eta_\xi(0,0))\) hold.

Letting \(n \to \infty\) in Equations S43 and S45, we obtain the relation

\[\sum_{i=1}^{N} \pi_i \bar{q}_i = \sum_{i,j \in S} \pi_i \tau_i t_{ij},\]
\[\sum_{i=1}^{N} \tilde{\pi}_i \bar{\tilde{q}}_i = \sum_{i,j \in S} \tilde{\pi}_i \tau_i t_{ij}.\] (S47)

The right-hand side of Equation S47 is denoted by \(B_0\), which is independent to \(\eta_\xi_{ij}\).

Let \(\mu_1 = \sum_{i=1}^{N} w_i / N\) (\(\tilde{\mu}_1 = \sum_{i=1}^{N} \tilde{w}_i / N\)) and \(\mu_2 = \sum_{i=1}^{N} w_i^2 / N\) (\(\tilde{\mu}_2 = \sum_{i=1}^{N} \tilde{w}_i^2 / N\)) denote the first and second moments of the network weighted degree distribution. Equation S47 is transformed to

\[\sum_{i=1}^{N} w_i^2 q_i = B_0 (N \mu_1)^2,\]
\[\sum_{i=1}^{N} \tilde{w}_i^2 \tilde{\tilde{q}}_i = B_0 (N \tilde{\mu}_1)^2.\] (S48)

Applying the mean-field approximation, we obtain

\[q_i = \frac{B_0 N \mu_1^2}{\mu_2},\]
\[\tilde{q}_i = \frac{B_0 N \tilde{\mu}_1^2}{\tilde{\mu}_2}.\] (S49)

In the case of uniform initialization, we obtain

\[\xi_i \xi_j = 0, \forall i \neq j,\]

and

\[\tilde{\xi} = \frac{1}{N}.\]
which gives $t_{ij} = \frac{1}{2}$ for all $i, j \in S$ and $B_0 = \frac{1}{2}$. Substituting these into Equations S43 and S45, we obtain the recurrence relation for uniform initialization, which is the same as Eq. (27) in Ref. [6].

As $p_{ii}^{(1)} = 0$ for all $i$, we approximate $\eta_{(n)}^{\xi}$ and $\eta_{(n)}^{\mu}$ as follows,

$$
\eta_{(1)}^{\xi} \approx \frac{B_0N\mu_1^2}{\mu_2} - \frac{1}{N\mu_1} \sum_{i=1}^{N} w_i t_{ii},
$$

$$
\eta_{(2)}^{\xi} \approx \frac{B_0N\mu_1^2}{\mu_2} - \frac{1}{N\mu_1} \left( \sum_{ij=1}^{N} w_i (p_{ij}^{(0)} + p_{ij}^{(1)}) t_{ij} \right),
$$

$$
\eta_{(3)}^{\xi} \approx \frac{B_0N\mu_1^2}{\mu_2} + \frac{B_0N\mu_1 \sum_{i=1}^{N} w_i p_{ii}^{(2)}}{\mu_2} - \frac{1}{N\mu_1} \left( \sum_{ij=1}^{N} w_i (p_{ij}^{(0)} + p_{ij}^{(1)} + p_{ij}^{(2)}) t_{ij} \right),
$$

$$
\eta_{(1)}^{\mu} \approx \frac{N\mu_1^2}{2\mu_2} - \frac{1}{2},
$$

$$\eta_{(2)}^{\mu} \approx \frac{N\mu_1^2}{2\mu_2} - 1,
$$

$$\eta_{(3)}^{\mu} \approx \frac{N\mu_1^2}{2\mu_2} + \frac{\sum_{i=1}^{N} w_i p_{ii}^{(2)}}{2\mu_2} - \frac{3}{2}.
$$

When $w_i \gg 1$ for all $i$, we have $p_{ij} \approx 1$ for all $i \neq j$. Considering $\tilde{p}_{ii}^{(1)} \neq 0$ for all $i$, we see that

$$
\eta_{(1,0)}^{\xi} \approx \eta_{(0,1)}^{\xi} \approx \frac{B_0N\tilde{\mu}_1^2}{\mu_2} - \frac{1}{N\tilde{\mu}_1} \sum_{i=1}^{N} \tilde{w}_i t_{ii},
$$

$$
\eta_{(2,0)}^{\xi} \approx \sum_{i=1}^{N} \pi_i (1 + \tilde{p}_{ii}^{(1)}) \frac{B_0N\tilde{\mu}_1^2}{\mu_2} - \frac{1}{N\tilde{\mu}_1} \left( \sum_{i,j=1}^{N} \tilde{w}_i (\tilde{p}_{ij}^{(0)} + \tilde{p}_{ij}^{(1)}) t_{ij} \right),
$$

$$
\eta_{(3,0)}^{\xi} \approx \eta_{(2,1)}^{\xi} \approx \sum_{i=1}^{N} \pi_i (1 + \tilde{p}_{ii}^{(1)} + \tilde{p}_{ii}^{(2)}) \frac{B_0N\tilde{\mu}_1^2}{\mu_2} - \frac{1}{N\tilde{\mu}_1} \left( \sum_{i,j=1}^{N} \tilde{w}_i (\tilde{p}_{ij}^{(0)} + \tilde{p}_{ij}^{(1)} + \tilde{p}_{ij}^{(2)}) t_{ij} \right),
$$

$$
\eta_{(1,0)}^{\mu} \approx \eta_{(0,1)}^{\mu} \approx \frac{N\tilde{\mu}_1^2}{2\mu_2} - \frac{1}{2},
$$

$$\eta_{(2,0)}^{\mu} \approx \sum_{i=1}^{N} \pi_i (1 + \tilde{p}_{ii}^{(1)}) \frac{N\tilde{\mu}_1^2}{2\mu_2} - 1,
$$

$$\eta_{(3,0)}^{\mu} \approx \eta_{(2,1)}^{\mu} \approx \sum_{i=1}^{N} \pi_i (1 + \tilde{p}_{ii}^{(1)} + \tilde{p}_{ii}^{(2)}) \frac{N\tilde{\mu}_1^2}{2\mu_2} - \frac{3}{2}.$$
Therefore, the fixation probabilities \((\rho_S^\xi)^*\) and \((\rho_S^\mu)^*\) can be approximated as

\[
(\rho_S^\xi)^* \approx \begin{cases} 
\xi + \frac{\delta}{N} \left( -c \left( \frac{N^2 \mu_1^2}{2 \mu_2} - 1 \right) + b \left( \frac{\Lambda \mu_1}{2 \mu_2} - 1 \right) \right) & \text{(DB updating)}, \\
\xi + \frac{\delta}{N} \left( -c \left( \frac{N^2 \mu_1^2}{2 \mu_2} - 1 \right) \right) & \text{(PC updating)}, \\
\xi + \frac{\delta}{N} \left( -c \left( \sum_{i=1}^N \tau_i \left( 1 + \tilde{p}_{ii}^{(1)} \right) \frac{N \mu_1^2}{2 \mu_2} - 1 \right) + b \left( \frac{\Lambda \tilde{\mu}_1}{2 \mu_2} - 1 \right) \right) & \text{(IM updating)}, 
\end{cases}
\]

and

\[
(\rho_S^\mu)^* \approx \begin{cases} 
\frac{1}{N} + \frac{\delta}{N} \left( -c \left( \frac{N^2 \mu_1^2}{2 \mu_2} - 1 \right) + b \left( \frac{\Lambda \mu_1}{2 \mu_2} - 1 \right) \right) & \text{(DB updating)}, \\
\frac{1}{N} + \frac{\delta}{N} \left( -c \left( \frac{N^2 \mu_1^2}{2 \mu_2} - 1 \right) - b \frac{1}{2} \right) & \text{(PC updating)}, \\
\frac{1}{N} + \frac{\delta}{N} \left( -c \left( \sum_{i=1}^N \tau_i \left( 1 + \tilde{p}_{ii}^{(1)} \right) \frac{N \mu_1^2}{2 \mu_2} - 1 \right) + b \left( \frac{\Lambda \tilde{\mu}_1}{2 \mu_2} - 1 \right) \right) & \text{(IM updating)}, 
\end{cases}
\]

where \(C_k = \sum_{i,j=1}^N w_i p_{ij}^{(k)} t_{ij} (\tilde{C}_k = \sum_{i,j=1}^N \tilde{w}_i \tilde{p}_{ij}^{(k)} t_{ij})\) for \(k = 0, 1, 2\), and \(\Lambda = \sum_{i=1}^N \tilde{w}_i (\tilde{p}_{ii}^{(1)} + \tilde{p}_{ii}^{(2)})\). The approximate critical benefit-to-cost ratio of \(S\) in the case of a specific configuration \(\xi\) and uniform initialization is given as

\[
\begin{align*}
\left( \frac{b}{c} \right)^*_{\xi} &\approx \begin{cases} 
\frac{B_0 N^2 \mu_1^3 - \mu_2 (C_0 + C_1)}{B_0 N \Lambda \mu_1^2 + \mu_2 (C_1 + C_2)} & \text{(DB updating)}, \\
\frac{\mu_2 C_0 - B_0 N^2 \mu_1^3}{\mu_2 C_1} & \text{(PC updating)}, \\
\frac{\sum_{i=1}^N \tau_i \left( 1 + \tilde{p}_{ii}^{(1)} \right) B_0 N^2 \mu_1^3 - \tilde{\mu}_2 (\tilde{C}_0 + \tilde{C}_1)}{B_0 N \Lambda \tilde{\mu}_1^2 + \tilde{\mu}_2 (C_1 + C_2)} & \text{(IM updating)}, 
\end{cases}
\end{align*}
\]

and

\[
\begin{align*}
\left( \frac{b}{c} \right)^*_{\mu} &\approx \begin{cases} 
\frac{N \mu_1^2 - 2 \mu_2}{\Lambda \mu_1 - 2 \mu_2} & \text{(DB updating)}, \\
\frac{\mu_2 - N \mu_1^2}{2 \mu_2} & \text{(PC updating)}, \\
\frac{\sum_{i=1}^N \tau_i \left( 1 + \tilde{p}_{ii}^{(1)} \right) N \mu_1^2 - 2 \mu_2}{\Lambda \tilde{\mu}_1 - 2 \tilde{\mu}_2} & \text{(IM updating)}. 
\end{cases}
\end{align*}
\]

All the notations in Equations S52-S55 are unrelated to \(\eta_{ij}^\xi\) and \(\eta_{ij}^\mu\), so that we can avoid solving the linear systems.
6 The monotonicity about $g$

Let $\rho_T^\mu(g)$ denote the fixation probability for cooperation in the second evolutionary process that the evolution lasts for $g$ generation in each snapshot (except the last one). Figures 4 and 5 in the main text shows that $(\rho_T^\mu(g))^*$ is monotonically increasing with respect to $g$. Here we give corresponding evidence under neutral drift. We first consider $T = \{G^{(1)}, G^{(2)}\}$. We assume that $G^{(1)} (G^{(2)})$ has $m$ $(m + \Delta m)$ nodes, the average connectivity is $k_1$ ($k_2$), and there are initially $j$ cooperators in $G^{(1)}$. After a one-step DB updating ($g = 1$), in the average sense, the probability of being $j - 1$ cooperators in $G^{(1)}$ is

$$\alpha_j = \frac{j}{m} \sum_{i=0}^{k_1} \frac{(m-j)(i-1)}{(m-k_1)} \frac{i}{k_1} = \frac{j(m-j)}{m(m-1)},$$

and the probability of being $j + 1$ cooperators is

$$\beta_j = \frac{m-j}{m} \sum_{i=0}^{k_1} \frac{(j)(m-j-1)}{(m-k_1)} \frac{i}{k_1} = \frac{j(m-j)}{m(m-1)},$$

and the probability of remaining $j$ cooperators is

$$\gamma_j = 1 - \alpha_j - \beta_j.$$

When entering $G^{(2)}$ with the above distribution of cooperators, and assuming that each cooperator adds $s$ edges on average, we obtain

$$(\rho_T^\mu(1))^* = \alpha_j \frac{(j-1)k_1 + (j-1)s}{(m+\Delta m)k_2} + \gamma_j \frac{jk_1 + js}{(m+\Delta m)k_2} + \beta_j \frac{(j+1)k_1 + (j+1)s}{(m+\Delta m)k_2} = \frac{jk_1}{(m+\Delta m)k_2}.$$ 

Then the condition for $(\rho_T^\mu(1))^* > (\rho_T^\mu(0))^* = \frac{j}{m+\Delta m}$ is

$$k_1 > k_2,$$

or the equivalent expression

$$\Delta K < \frac{k_1\Delta m}{2},$$

which is similar to Equation 6b in the main text. Note that Equations S56 and S57 are independent to the number of cooperators $j$.

Based on the above analysis, after a one-step update, a single case of $j$ cooperators will be split into three cases $(j - 1, j, j + 1)$ with a certain distribution. When Equation S57 holds, a single-step update can increase the fixation probability. When the number of rounds $g$ increases, the number of the update also increases, so it is more likely to promote the evolution of cooperation when the parameter $g$ is larger. The ability of the promote reaches the
maximum when the parameter $g = \infty$. This explains why the fixation probability increases monotonically with $g$.

7 Algorithms for constructing sequential temporal networks

7.1 Synthetic networks

We provide a detailed construction of the sequential temporal networks based on square lattices, random regular graphs, Barabási-Albert scale-free networks, and scale-free networks with initial attractiveness mentioned in the main text. Supplementary Figures 1 and 2 show the construction for a square lattice and a Barabási-Albert scale-free network, respectively. The network construction process of scale-free networks with the initial attractiveness is quite similar to Barabási-Albert scale-free networks.

For random regular graphs, we provide a general construction algorithm, which can also be used on square lattices, scale-free networks, and other network structures.

**Algorithm 1 Construction of sequential temporal networks**

**Input:** A static network topology $\mathcal{S}$.

**Output:** A sequential temporal network $\mathcal{T}$.

1: Randomly select a node and $m_0$ its neighbours to form an initial snapshot;
2: while $\#$(active nodes) $< N$ do
3: Randomly select a node, which is a neighbour of an active node, to be activated;

This algorithm ensures that each snapshot is connected.

7.2 Empirical networks

For empirical data analysis, we use four datasets collected by the SocioPatterns collaboration (http://www.sociopatterns.org). The dataset of the infectious dynamic contact is divided according to the date, and we select the first three days (2009.04.28-2009.04.30) for analysis. For the other three datasets, we use the full data.

In the datasets, a contact event is formatted as a triplet $(t, i, j)$, which represents an interaction between two individuals $i$ and $j$ occurring at time $t$. Based on the triplets, we construct the corresponding static and temporal networks. In the first step, we aggregate all contacts into a single undirected graph denoted by $\mathcal{S}_{\text{agg}}$. If the graph $\mathcal{S}_{\text{agg}}$ is not connected, we select the giant component of $\mathcal{S}_{\text{agg}}$ as a new static network. Subsequently, we determine the activity of individuals according to the contact events. For each individual, we find the first communication time $t_1$ and set $a_t^{(j)} = 1$ for all $j \geq t_1$. Then, we have a vector set $\mathcal{A} = \{a^{(1)}, \ldots, a^{(L)}\}$, and
the sequential temporal network $\mathcal{T} = \{G^{(1)}, ..., G^{(L)}\}$ is generated by $S_{agg}$ and $A$. Since each snapshot $G^{(i)}$ may be disconnected, we select all connected snapshots to form the sequential temporal network $\mathcal{T} = \{G^{(s)}, ..., G_{agg}\}$.

References


