Supplementary information
A scaling law of multilevel evolution: how the balance between within- and among-collective evolution is determined

Nobuto Takeuchi*[1,4], Namiko Mitarai[2,4], and Kunihiko Kaneko[3,4]

[1]School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
[2]The Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, Copenhagen, 2100-DK, Denmark
[3]Graduate School of Arts and Sciences, University of Tokyo, Komaba 3-8-1, Meguro-ku, Tokyo 153-8902, Japan
[4]Research Center for Complex Systems Biology, Universal Biology Institute, University of Tokyo, Komaba 3-8-1, Meguro-ku, Tokyo 153-8902, Japan

1 Derivation of Eq. (3)

In this section, we derive Eq. (3) of the main text, which is redisplayed below:

$$\mathbb{E}[\Delta\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle] = \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\left\{\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, \langle k_{i\tilde{j}}\rangle\right] + \mathrm{ave}_{\tilde{i}}\left[\mathrm{cov}_{i\tilde{j}}\left[w_{ij}, k_{ij}\right]\right]\right\},$$

where the symbols are defined as follows:

$$\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle := \frac{1}{M}\sum_{i=1}^{L} n_i \langle k_{i\tilde{j}}\rangle, \tag{S1}$$

*Corresponding author. E-mail: nobuto.takeuchi@auckland.ac.nz

where M is the total number of replicators, n_i is the number of replicators in collective i, L is the number of collectives, and

$$
\begin{aligned}
&\langle k_{i\tilde{j}}\rangle := \frac{1}{n_i}\sum_{j=1}^{n_i} k_{ij}, \\
&\langle\langle w_{\tilde{i}\tilde{j}}\rangle\rangle := \frac{1}{M}\sum_{i=1}^{L} n_i \langle w_{i\tilde{j}}\rangle, \\
&\langle w_{i\tilde{j}}\rangle := \frac{1}{n_i}\sum_{j=1}^{n_i} w_{ij}, \\
&\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, \langle k_{i\tilde{j}}\rangle\right] := \frac{1}{M}\sum_{i=1}^{L} n_i \left(\langle w_{i\tilde{j}}\rangle - \langle\langle w_{\tilde{i}\tilde{j}}\rangle\rangle\right)\left(\langle k_{i\tilde{j}}\rangle - \langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle\right), \\
&\mathrm{cov}_{i\tilde{j}}\left[w_{ij}, k_{ij}\right] := \frac{1}{n_i}\sum_{j=1}^{n_i} \left(w_{ij} - \langle w_{i\tilde{j}}\rangle\right)\left(k_{ij} - \langle k_{i\tilde{j}}\rangle\right), \\
&\mathrm{ave}_{\tilde{i}}\left[\mathrm{cov}_{i\tilde{j}}\left[w_{ij}, k_{ij}\right]\right] := \frac{1}{M}\sum_{i=1}^{L} n_i\, \mathrm{cov}_{i\tilde{j}}\left[w_{ij}, k_{ij}\right].
\end{aligned}
\tag{S2}
$$

In each generation, a replicator is sampled M times with replacement from replicators of the previous generation with probabilities proportional to fitness w_{ij}, as in the Wright-Fisher process (see the main text under "Model"). To express $\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle$ in the next generation, we introduce the following symbols. Let I_l be the index of the collective to which the lth sampled replicator belongs ($l \in \{1, 2, \cdots, M\}$), J_l be the index of the sampled replicator within collective I_l, and $P(I_l = i, J_l = j)$ be the probability that replicator j in collective i is sampled. By the definition of the Wright-Fisher process,

$$P(I_l = i,\ J_l = j) = \frac{w_{ij}}{M\langle\langle w_{\tilde{i}\tilde{j}}\rangle\rangle}. \tag{S3}$$

Moreover, let $\epsilon_{I_l J_l}$ be the effect of mutation and $P(\epsilon_{I_l J_l})$ be its probability distribution function ($\epsilon_{I_l J_l}$ takes a value of 0 with a probability $1 - m$ or a value sampled from a Gaussian distribution with mean 0 and variance σ with a probability m). Finally, let $\mathbb{E}[x]$ denote the expected value of x after one iteration of the Wright-Fisher process; e.g.,

$$\mathbb{E}[k_{I_l J_l} + \epsilon_{I_l J_l}] = \sum_{i=1}^{L}\sum_{j=1}^{n_i} P(I_l = i, J_l = j) \int dP(\epsilon_{I_l J_l})(k_{ij} + \epsilon_{I_l J_l}). \tag{S4}$$

Using these definitions, we can express the expected change of $\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle$ per generation, denoted by $\mathbb{E}[\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle]$, as follows:

$$\mathbb{E}[\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle] = \mathbb{E}\left[M^{-1}\sum_{l=1}^{M}(k_{I_l J_l} + \epsilon_{I_l J_l})\right] - \langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle. \tag{S5}$$

Since I_l and J_l are independent and identically distributed for different values of l, we can remove the summation in the above equation to obtain

$$
\begin{aligned}
\mathbb{E}[\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle] &= \mathbb{E}\left[k_{IJ} + \epsilon_{IJ}\right] - \langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle \\
&= \mathbb{E}\left[k_{IJ}\right] - \langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle,
\end{aligned}
\tag{S6}
$$

where we used the fact that $\mathbb{E}\left[\epsilon_{IJ}\right] = 0$ and omitted subscript l.

The first term on the RHS of Eq. (S6) can be calculated as follows:

$$
\begin{aligned}
\mathbb{E}\left[k_{IJ}\right] &= \sum_{i=1}^{L}\sum_{j=1}^{n_i} P(I=i, J=j) k_{ij} \\
&= \sum_{i=1}^{L}\sum_{j=1}^{n_i} \frac{w_{ij}}{M\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle} k_{ij} \\
&= \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1} \frac{1}{M}\sum_{i=1}^{L}\sum_{j=1}^{n_i} w_{ij}k_{ij} \\
&= \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1} \frac{1}{M}\sum_{i=1}^{L} n_i \frac{1}{n_i}\sum_{j=1}^{n_i} w_{ij}k_{ij} \\
&= \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1} \frac{1}{M}\sum_{i=1}^{L} n_i \left(\langle w_{i\tilde{j}}\rangle\langle k_{i\tilde{j}}\rangle + \mathrm{cov}_{i\tilde{j}}\left[w_{ij}, k_{ij}\right]\right) \\
&= \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1} \left\{ \frac{1}{M}\Big(\sum_{i=1}^{L} n_i \langle w_{i\tilde{j}}\rangle\langle k_{i\tilde{j}}\rangle\Big) + \mathrm{ave}_{\tilde{i}}\left[\mathrm{cov}_{i\tilde{j}}\left[w_{ij}, k_{ij}\right]\right] \right\} \\
&= \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1} \left(\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, \langle k_{i\tilde{j}}\rangle\right] + \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle + \mathrm{ave}_{\tilde{i}}\left[\mathrm{cov}_{i\tilde{j}}\left[w_{ij}, k_{ij}\right]\right]\right) \\
&= \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle + \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1} \left\{\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, \langle k_{i\tilde{j}}\rangle\right] + \mathrm{ave}_{\tilde{i}}\left[\mathrm{cov}_{i\tilde{j}}\left[w_{ij}, k_{ij}\right]\right]\right\}.
\end{aligned} \tag{S7}
$$

Substituting Eq. (S7) into Eq. (S6), we obtain Eq. (3).

2 Derivation of Eq. (4)

In this section, we derive Eq. (4) of the main text, which is redisplayed below:

$$
\mathbb{E}[\Delta\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle] = s_{\mathrm{a}}v_{\mathrm{a}} - s_{\mathrm{w}}v_{\mathrm{w}} + O(s_{\mathrm{w}}^2) + O(s_{\mathrm{a}}^2),
$$

where the symbols are defined as follows:

$$
\begin{aligned}
v_{\mathrm{a}} &:= \mathrm{cov}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle, \langle k_{i\tilde{j}}\rangle\right] = \frac{1}{M}\sum_{i=1}^{L} n_i \left(\langle k_{i\tilde{j}}\rangle - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\right)^2 \\
v_{\mathrm{w}i} &:= \mathrm{cov}_{i\tilde{j}}\left[k_{ij}, k_{ij}\right] = \frac{1}{n_i}\sum_{j=1}^{n_i} \left(k_{ij} - \langle k_{i\tilde{j}}\rangle\right)^2, \\
v_{\mathrm{w}} &:= \mathrm{ave}_{\tilde{i}}\left[v_{\mathrm{w}i}\right] = \frac{1}{M}\sum_{i=1}^{L} n_i\, v_{\mathrm{w}i}.
\end{aligned} \tag{S8}
$$

Equation (4) is obtained by expanding $\langle w_{i\tilde{j}}\rangle$ and w_{ij} in Eq. (3), i.e.,

$$
\Delta\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle = \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, \langle k_{i\tilde{j}}\rangle\right] + \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{ave}_{\tilde{i}}\left[\mathrm{cov}_{i\tilde{j}}\left[w_{ij}, k_{ij}\right]\right],
$$

as Taylor series around $\langle k_{i\tilde{j}}\rangle = \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle$ and $k_{ij} = \langle k_{i\tilde{j}}\rangle$, respectively.

First, we obtain the first term of Eq. (4), which stems from the first term of Eq. (3). We assume that $\langle w_{i\tilde{\jmath}}\rangle$ is an analytic function of $\langle k_{i\tilde{\jmath}}\rangle$ and that $\langle w_{i\tilde{\jmath}}\rangle = \langle\!\langle w_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle$ for $\langle k_{i\tilde{\jmath}}\rangle = \langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle$. Expanding $\langle w_{i\tilde{\jmath}}\rangle$ around $\langle k_{i\tilde{\jmath}}\rangle = \langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle$, we obtain

$$\langle w_{i\tilde{\jmath}}\rangle = \langle\!\langle w_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle + \frac{\partial\langle w_{i\tilde{\jmath}}\rangle}{\partial\langle k_{i\tilde{\jmath}}\rangle}\left(\langle k_{i\tilde{\jmath}}\rangle - \langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle\right) + \frac{\partial^2\langle w_{i\tilde{\jmath}}\rangle}{\partial\langle k_{i\tilde{\jmath}}\rangle^2}\left(\langle k_{i\tilde{\jmath}}\rangle - \langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle\right)^2 + \cdots. \tag{S9}$$

Dividing both sides by $\langle\!\langle w_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle$, we obtain

$$\frac{\langle w_{i\tilde{\jmath}}\rangle}{\langle\!\langle w_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle} = 1 + \frac{1}{\langle\!\langle w_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle}\frac{\partial\langle w_{i\tilde{\jmath}}\rangle}{\partial\langle k_{i\tilde{\jmath}}\rangle}\left(\langle k_{i\tilde{\jmath}}\rangle - \langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle\right) + \frac{1}{\langle\!\langle w_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle}\frac{\partial^2\langle w_{i\tilde{\jmath}}\rangle}{\partial\langle k_{i\tilde{\jmath}}\rangle^2}\left(\langle k_{i\tilde{\jmath}}\rangle - \langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle\right)^2 + \cdots. \tag{S10}$$

By the definition of selection strength (see the main text under "Model"),

$$\begin{aligned}\frac{1}{\langle\!\langle w_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle}\frac{\partial\langle w_{i\tilde{\jmath}}\rangle}{\partial\langle k_{i\tilde{\jmath}}\rangle}\bigg|_{\langle k_{i\tilde{\jmath}}\rangle=\langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle} &= \left(\frac{1}{\langle w_{i\tilde{\jmath}}\rangle}\frac{\partial\langle w_{i\tilde{\jmath}}\rangle}{\partial\langle k_{i\tilde{\jmath}}\rangle}\right)\bigg|_{\langle k_{i\tilde{\jmath}}\rangle=\langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle} \\ &:= s_{\mathrm{a}}\end{aligned} \tag{S11}$$

By mathematical induction, it can be shown that

$$\frac{1}{\langle w_{i\tilde{\jmath}}\rangle}\frac{\partial^{l+1}\langle w_{i\tilde{\jmath}}\rangle}{\partial\langle k_{i\tilde{\jmath}}\rangle^{l+1}} = \left(\frac{\partial}{\partial\langle k_{i\tilde{\jmath}}\rangle} + \frac{1}{\langle w_{i\tilde{\jmath}}\rangle}\frac{\partial\langle w_{i\tilde{\jmath}}\rangle}{\partial\langle k_{i\tilde{\jmath}}\rangle}\right)\frac{1}{\langle w_{i\tilde{\jmath}}\rangle}\frac{\partial^{l}\langle w_{i\tilde{\jmath}}\rangle}{\partial\langle k_{i\tilde{\jmath}}\rangle^{l}} \tag{S12}$$

for $l \in \{1, 2, 3, \cdots\}$. Since it is assumed that $\partial s_{\mathrm{a}}/\partial\langle k_{i\tilde{\jmath}}\rangle = 0$ (see the main text under "Model"), the above equation implies that

$$\left(\frac{1}{\langle w_{i\tilde{\jmath}}\rangle}\frac{\partial^{l}\langle w_{i\tilde{\jmath}}\rangle}{\partial\langle k_{i\tilde{\jmath}}\rangle^{l}}\right)\bigg|_{\langle k_{i\tilde{\jmath}}\rangle=\langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle} = s_{\mathrm{a}}^{l}. \tag{S13}$$

Given the above equation, Eq. (S10) implies that

$$\frac{\langle w_{i\tilde{\jmath}}\rangle}{\langle\!\langle w_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle} = 1 + s_{\mathrm{a}}\left(\langle k_{i\tilde{\jmath}}\rangle - \langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle\right) + O(s_{\mathrm{a}}^2). \tag{S14}$$

Therefore,

$$\begin{aligned}\langle\!\langle w_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{\imath}}\left[\langle w_{i\tilde{\jmath}}\rangle, \langle k_{i\tilde{\jmath}}\rangle\right] &= \mathrm{cov}_{\tilde{\imath}}\left[1 + s_{\mathrm{a}}\left(\langle k_{i\tilde{\jmath}}\rangle - \langle\!\langle k_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle\right) + O(s_{\mathrm{a}}^2), \langle k_{i\tilde{\jmath}}\rangle\right] \\ &= s_{\mathrm{a}}v_{\mathrm{a}} + O(s_{\mathrm{a}}^2)\end{aligned} \tag{S15}$$

Second, we obtain the second term of Eq. (4), which stems from the second term of Eq. (3). Using the same method as above, we can show that

$$\langle w_{i\tilde{\jmath}}\rangle^{-1}\mathrm{cov}_{i\tilde{\jmath}}\left[w_{ij}, k_{ij}\right] = -s_{\mathrm{w}}v_{\mathrm{w}i} + O(s_{\mathrm{w}}^2). \tag{S16}$$

We assume that $\langle w_{i\tilde{\jmath}}\rangle$ and $v_{\mathrm{w}i}$ are statistically uncorrelated as i varies (this is equivalent to assuming that $\langle k_{i\tilde{\jmath}}\rangle$ and $v_{\mathrm{w}i}$ are uncorrelated). Under this assumption,

$$\begin{aligned}\mathrm{ave}_{\tilde{\imath}}\left[\mathrm{cov}_{i\tilde{\jmath}}\left[w_{ij}, k_{ij}\right]\right] &= \mathrm{ave}_{\tilde{\imath}}\left[-\langle w_{i\tilde{\jmath}}\rangle\, s_{\mathrm{w}}\, v_{\mathrm{w}i} + O(s_{\mathrm{w}}^2)\right] \\ &= -\mathrm{ave}_{\tilde{\imath}}\left[\langle w_{i\tilde{\jmath}}\rangle\right] s_{\mathrm{w}}\,\mathrm{ave}_{\tilde{\imath}}\left[v_{\mathrm{w}i}\right] + O(s_{\mathrm{w}}^2) \\ &= -\langle\!\langle w_{\tilde{\imath}\tilde{\jmath}}\rangle\!\rangle s_{\mathrm{w}}v_{\mathrm{w}} + O(s_{\mathrm{w}}^2).\end{aligned} \tag{S17}$$

Substituting Eqs. (S15) and (S17) into Eq. (3), we obtain Eq. (4).

3 Derivation of Eq. (8)

In this section, we derive Eq. (8), which is redisplayed below:

$$\begin{aligned}\mathbb{E}\left[v'_{\mathrm{w}}\right] &= \left(1-\beta N^{-1}\right)\left[v_{\mathrm{w}} + m\sigma - s_{\mathrm{w}}c_{\mathrm{w}} + O(s_{\mathrm{w}}^2)\right] \\ \mathbb{E}\left[v'_{\mathrm{a}}\right] &= \left(1-M^{-1}\right)\left[v_{\mathrm{a}} + s_{\mathrm{a}}c_{\mathrm{a}} + O\left((s_{\mathrm{w}}+s_{\mathrm{a}})^2\right)\right] \\ &\quad + \left(\beta N^{-1} - M^{-1}\right)\left[v_{\mathrm{w}} + m\sigma - s_{\mathrm{w}}c_{\mathrm{w}} + O(s_{\mathrm{w}}^2)\right],\end{aligned}$$

3.1 Calculation of $\mathbb{E}[v'_{\mathrm{w}}]$

To calculate $\mathbb{E}[v'_{\mathrm{w}}]$, we introduce the following symbols. Let n'_i be the number of replicators in collective i after one iteration of the Wright-Fisher process. Note that n'_i is a random variable and can be expressed as

$$n'_i = \sum_{l=1}^{M} \delta_{I_l i} \tag{S18}$$

where $\delta_{I_l i}$ is the Kronecker delta (i.e., $\delta_{I_l i} = 1$ if $I_l = i$, and $\delta_{I_l i} = 0$ otherwise). Moreover, let $\langle k_{iJ_{\tilde{l}}}\rangle$ and $\langle \epsilon_{iJ_{\tilde{l}}}\rangle$ be the sample mean of $k_{I_l J_l}$ and $\epsilon_{I_l J_l}$ within collective i:

$$\begin{aligned}\langle k_{iJ_{\tilde{l}}}\rangle &:= \frac{1}{n'_i}\sum_{l=1}^{n'_i} k_{iJ_l} \\ \langle \epsilon_{iJ_{\tilde{l}}}\rangle &:= \frac{1}{n'_i}\sum_{l=1}^{n'_i} \epsilon_{I_l J_l},\end{aligned} \tag{S19}$$

which are defined to be zero when $n'_i = 0$. The probability that $J_l = j$ given $I_l = i$ is

$$\begin{aligned}P(J=j|I=i) &= \frac{P(I=i, J=i)}{P(I=i)} \\ &= \frac{P(I=i,J=i)}{\sum_{j=1}^{n_i} P(I=i,J=i)} \\ &= \frac{\frac{w_{ij}}{M\langle\langle w_{\tilde{i}\tilde{j}}\rangle\rangle}}{\sum_{j=1}^{n_i}\frac{w_{ij}}{M\langle\langle w_{\tilde{i}\tilde{j}}\rangle\rangle}} \\ &= \frac{\frac{w_{ij}}{M\langle\langle w_{\tilde{i}\tilde{j}}\rangle\rangle}}{\frac{n_i\langle w_{i\tilde{j}}\rangle}{M\langle\langle w_{\tilde{i}\tilde{j}}\rangle\rangle}} \\ &= \frac{w_{ij}}{n_i\langle w_{i\tilde{j}}\rangle},\end{aligned} \tag{S20}$$

where Eq. (S3) is used.

Using the symbols defined above, we can express the sample variance of k_{ij} within collective i in the next generation as

$$v'_{\mathrm{w}i} := \frac{1}{n'_i}\sum_{l=1}^{n'_i}\left(k_{iJ_l} + \epsilon_{iJ_l} - \langle k_{iJ_{\tilde{l}}}\rangle - \langle \epsilon_{iJ_{\tilde{l}}}\rangle\right)^2, \tag{S21}$$

which is defined to be zero when $n'_i = 0$. Using $v'_{\mathrm{w}i}$, we can express $\mathbb{E}[v'_{\mathrm{w}}]$ as follows:

$$\mathbb{E}[v'_{\mathrm{w}}] = \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L} n'_i v'_{\mathrm{w}i}\right]. \tag{S22}$$

In the last equation, we can separate k_{iJ_l} and ϵ_{iJ_l} as follows:

$$\begin{aligned}
\mathbb{E}[v'_{\mathrm{w}}] &= \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L} n'_i \frac{1}{n'_i}\sum_{l=1}^{n'_i}\left(k_{iJ_l} + \epsilon_{iJ_l} - \langle k_{iJ_{\tilde{l}}}\rangle - \langle \epsilon_{iJ_{\tilde{l}}}\rangle\right)^2\right] \\
&= \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n'_i}\left\{\left(k_{iJ_l} - \langle k_{iJ_{\tilde{l}}}\rangle\right)^2 + \left(\epsilon_{iJ_l} - \langle \epsilon_{iJ_{\tilde{l}}}\rangle\right)^2 + 2\left(k_{iJ_l} - \langle k_{iJ_{\tilde{l}}}\rangle\right)\left(\epsilon_{iJ_l} - \langle \epsilon_{iJ_{\tilde{l}}}\rangle\right)\right\}\right] \\
&= \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n'_i}\left(k_{iJ_l} - \langle k_{iJ_{\tilde{l}}}\rangle\right)^2\right] + \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n'_i}\left(\epsilon_{iJ_l} - \langle \epsilon_{iJ_{\tilde{l}}}\rangle\right)^2\right] \\
&\quad + 2\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n'_i}\left(k_{iJ_l} - \langle k_{iJ_{\tilde{l}}}\rangle\right)\left(\epsilon_{iJ_l} - \langle \epsilon_{iJ_{\tilde{l}}}\rangle\right)\right].
\end{aligned} \tag{S23}$$

The last term of the final line of Eq. (S23) can be shown to be zero, as follows. With the Kronecker delta $\delta_{I_l i}$, this term can be calculated as

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n'_i}\left(k_{iJ_l} - \langle k_{iJ_{\tilde{l}}}\rangle\right)\left(\epsilon_{iJ_l} - \langle \epsilon_{iJ_{\tilde{l}}}\rangle\right)\right] \\
&= \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{M}\delta_{I_l i}\left(k_{I_l J_l} - \langle k_{iJ_{\tilde{l}}}\rangle\right)\left(\epsilon_{iJ_l} - \langle \epsilon_{iJ_{\tilde{l}}}\rangle\right)\right] \\
&= \frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{M}\mathbb{E}\left[\delta_{I_l i}\left(k_{I_l J_l} - \langle k_{iJ_{\tilde{l}}}\rangle\right)\right]\mathbb{E}\left[\left(\epsilon_{iJ_l} - \langle \epsilon_{iJ_{\tilde{l}}}\rangle\right)\right] \\
&= 0,
\end{aligned} \tag{S24}$$

where we used the fact that k_{iJ_l} and ϵ_{iJ_l} are independent of each other. Therefore,

$$\mathbb{E}[v'_{\mathrm{w}}] = \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n'_i}\left(k_{iJ_l} - \langle k_{iJ_{\tilde{l}}}\rangle\right)^2\right] + \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n'_i}\left(\epsilon_{iJ_l} - \langle \epsilon_{iJ_{\tilde{l}}}\rangle\right)^2\right]. \tag{S25}$$

To calculate the first term of Eq. (S25), we define within-collective conditional expectation as follows:

$$\mathbb{E}_{J|I=i}[x_{IJ}] := \sum_{j=1}^{n_i} P(J=j|I=i)x_{ij}. \tag{S26}$$

Using Eq. (S26), we can transform the first term of Eq. (S25) as follows:

$$
\begin{aligned}
&\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n'_i}\left(k_{iJ_l}-\langle k_{iJ_{\tilde{l}}}\rangle\right)^2\right]\\
&=\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n'_i}\left(k_{iJ_l}-\mathbb{E}_{J|I=i}[k_{IJ}]+\mathbb{E}_{J|I=i}[k_{IJ}]-\langle k_{iJ_{\tilde{l}}}\rangle\right)^2\right]\\
&=\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\left\{\sum_{l=1}^{n'_i}\left(k_{iJ_l}-\mathbb{E}_{J|I=i}[k_{IJ}]\right)^2+\sum_{l=1}^{n'_i}\left(\mathbb{E}_{J|I=i}[k_{IJ}]-\langle k_{iJ_{\tilde{l}}}\rangle\right)^2\right.\right.\\
&\qquad\left.\left.+2\sum_{l=1}^{n'_i}\left(k_{iJ_l}-\mathbb{E}_{J|I=i}[k_{IJ}]\right)\left(\mathbb{E}_{J|I=i}[k_{IJ}]-\langle k_{iJ_{\tilde{l}}}\rangle\right)\right\}\right]\\
&=\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\left\{\sum_{l=1}^{n'_i}\left(k_{iJ_l}-\mathbb{E}_{J|I=i}[k_{IJ}]\right)^2+n'_i\left(\mathbb{E}_{J|I=i}[k_{IJ}]-\langle k_{iJ_{\tilde{l}}}\rangle\right)^2\right.\right.\\
&\qquad\left.\left.+2\left(\mathbb{E}_{J|I=i}[k_{IJ}]-\langle k_{iJ_{\tilde{l}}}\rangle\right)\sum_{l=1}^{n'_i}\left(k_{iJ_l}-\mathbb{E}_{J|I=i}[k_{IJ}]\right)\right\}\right]\\
&=\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\left\{\sum_{l=1}^{n'_i}\left(k_{iJ_l}-\mathbb{E}_{J|I=i}[k_{IJ}]\right)^2+n'_i\left(\mathbb{E}_{J|I=i}[k_{IJ}]-\langle k_{iJ_{\tilde{l}}}\rangle\right)^2\right.\right.\\
&\qquad\left.\left.+2\left(\mathbb{E}_{J|I=i}[k_{IJ}]-\langle k_{iJ_{\tilde{l}}}\rangle\right)n'_i\left(\langle k_{iJ_{\tilde{l}}}\rangle-\mathbb{E}_{J|I=i}[k_{IJ}]\right)\right\}\right]\\
&=\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\left\{\sum_{l=1}^{n'_i}\left(k_{iJ_l}-\mathbb{E}_{J|I=i}[k_{IJ}]\right)^2-n'_i\left(\langle k_{iJ_{\tilde{l}}}\rangle-\mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right\}\right]\\
&=\frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[\sum_{l=1}^{n'_i}\left(k_{iJ_l}-\mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right]-\frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[n'_i\left(\langle k_{iJ_{\tilde{l}}}\rangle-\mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right].
\end{aligned}
\tag{S27}
$$

The first term in the last line of Eq. (S27) is calculated as follows:

$$
\begin{aligned}
&\frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[\sum_{l=1}^{n'_i}\left(k_{iJ_l}-\mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right]\\
=&\frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[\sum_{l=1}^{n'_i}\left(k_{iJ_l}^2-2\mathbb{E}_{J|I=i}[k_{IJ}]k_{iJ_l}+\mathbb{E}_{J|I=i}[k_{IJ}]^2\right)\right]\\
=&\frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[\sum_{l=1}^{n'_i}k_{iJ_l}^2-2\mathbb{E}_{J|I=i}[k_{IJ}]\sum_{l=1}^{n'_i}k_{iJ_l}+n'_i\mathbb{E}_{J|I=i}[k_{IJ}]^2\right]\\
=&\frac{1}{M}\sum_{i=1}^{L}\left\{\mathbb{E}\left[\sum_{l=1}^{n'_i}k_{iJ_l}^2\right]-2\mathbb{E}_{J|I=i}[k_{IJ}]\mathbb{E}\left[\sum_{l=1}^{n'_i}k_{iJ_l}\right]+\mathbb{E}[n'_i]\mathbb{E}_{J|I=i}[k_{IJ}]^2\right\}.
\end{aligned}
\tag{S28}
$$

Using the Kronecker delta $\delta_{I_l i}$, we can show that

$$\begin{aligned}
\mathbb{E}\Bigg[\sum_{l=1}^{n_i'} k_{iJ_l}\Bigg] &= \mathbb{E}\Bigg[\sum_{l=1}^{M} \delta_{I_l i} k_{I_l J_l}\Bigg] \\
&= \sum_{l=1}^{M} \mathbb{E}\left[\delta_{I_l i} k_{I_l J_l}\right] \\
&= M\mathbb{E}\left[\delta_{I i} k_{IJ}\right] \\
&= M \sum_{j=1}^{n_i} P(I = i, J = j) k_{ij} \\
&= MP(I = i) \sum_{j=1}^{n_i} P(J = j | I = i) k_{ij} \\
&= M \frac{n_i \langle w_{i\tilde{j}} \rangle}{M \langle\langle w_{\tilde{i}\tilde{j}} \rangle\rangle} \mathbb{E}_{J|I=i}[k_{IJ}] \\
&= \frac{n_i \langle w_{i\tilde{j}} \rangle}{\langle\langle w_{\tilde{i}\tilde{j}} \rangle\rangle} \mathbb{E}_{J|I=i}[k_{IJ}].
\end{aligned} \tag{S29}$$

Likewise, we can show that

$$\begin{aligned}
\mathbb{E}\Bigg[\sum_{l=1}^{n_i'} k_{iJ_l}^2\Bigg] &= \mathbb{E}\Bigg[\sum_{l=1}^{M} \delta_{I_l i} k_{I_l J_l}^2\Bigg] \\
&= \sum_{l=1}^{M} \mathbb{E}\left[\delta_{I_l i} k_{I_l J_l}^2\right] \\
&= M\mathbb{E}\left[\delta_{I i} k_{IJ}^2\right] \\
&= M \sum_{j=1}^{n_i} P(I = i, J = j) k_{ij}^2 \\
&= MP(I = i) \sum_{j=1}^{n_i} P(J = j | I = i) k_{ij}^2 \\
&= \frac{n_i \langle w_{i\tilde{j}} \rangle}{\langle\langle w_{\tilde{i}\tilde{j}} \rangle\rangle} \mathbb{E}_{J|I=i}[k_{IJ}^2].
\end{aligned} \tag{S30}$$

Also, we can show that

$$\begin{aligned}\mathbb{E}[n_i'] &= \mathbb{E}\left[\sum_{l=1}^{M}\delta_{I_l i}\right]\\ &= \sum_{l=1}^{M}\mathbb{E}[\delta_{I_l i}]\\ &= M\mathbb{E}[\delta_{Ii}]\\ &= M\sum_{j=1}^{n_i}P(I=i,J=j)\\ &= MP(I=i)\\ &= \frac{n_i\langle w_{i\tilde{j}}\rangle}{\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}.\end{aligned} \tag{S31}$$

Using the above results, we can transform the last line of Eq. (S28) as follows:

$$\begin{aligned}&\frac{1}{M}\sum_{i=1}^{L}\left\{\mathbb{E}\left[\sum_{l=1}^{n_i'}k_{iJ_l}^2\right]-2\mathbb{E}_{J|I=i}[k_{IJ}]\mathbb{E}\left[\sum_{l=1}^{n_i'}k_{iJ_l}\right]+\mathbb{E}[n_i']\mathbb{E}_{J|I=i}[k_{IJ}]^2\right\}\\ =&\frac{1}{M}\sum_{i=1}^{L}\left\{\frac{n_i\langle w_{i\tilde{j}}\rangle}{\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\mathbb{E}_{J|I=i}[k_{IJ}^2]-2\mathbb{E}_{J|I=i}[k_{IJ}]\frac{n_i\langle w_{i\tilde{j}}\rangle}{\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\mathbb{E}_{J|I=i}[k_{IJ}]+\frac{n_i\langle w_{i\tilde{j}}\rangle}{\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\mathbb{E}_{J|I=i}[k_{IJ}]^2\right\}\\ =&\frac{1}{M}\sum_{i=1}^{L}\left\{\frac{n_i\langle w_{i\tilde{j}}\rangle}{\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\mathbb{E}_{J|I=i}[k_{IJ}^2]-\frac{n_i\langle w_{i\tilde{j}}\rangle}{\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\mathbb{E}_{J|I=i}[k_{IJ}]^2\right\}\\ =&\frac{1}{M}\sum_{i=1}^{L}\frac{n_i\langle w_{i\tilde{j}}\rangle}{\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\left\{\mathbb{E}_{J|I=i}[k_{IJ}^2]-\mathbb{E}_{J|I=i}[k_{IJ}]^2\right\}.\end{aligned} \tag{S32}$$

The conditional expectation in the last line of Eq. (S32) can be calculated as follows:

$$\begin{aligned}&\mathbb{E}_{J|I=i}[k_{IJ}^2]-\mathbb{E}_{J|I=i}[k_{IJ}]^2\\ &=\sum_{j=1}^{n_i}P(J=j|i=i)k_{ij}^2-\left(\sum_{j=1}^{n_i}P(J=j|i=i)k_{ij}\right)^2\\ &=\sum_{j=1}^{n_i}\frac{w_{ij}}{n_i\langle w_{i\tilde{j}}\rangle}k_{ij}^2-\left(\sum_{j=1}^{n_i}\frac{w_{ij}}{n_i\langle w_{i\tilde{j}}\rangle}k_{ij}\right)^2\\ &=\langle w_{i\tilde{j}}\rangle^{-1}\frac{1}{n_i}\sum_{j=1}^{n_i}w_{ij}k_{ij}^2-\left(\langle w_{i\tilde{j}}\rangle^{-1}\frac{1}{n_i}\sum_{j=1}^{n_i}w_{ij}k_{ij}\right)^2\\ &=\langle w_{i\tilde{j}}\rangle^{-1}\left(\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}^2]+\langle w_{i\tilde{j}}\rangle\langle k_{i\tilde{j}}^2\rangle\right)-\left\{\langle w_{i\tilde{j}}\rangle^{-1}\left(\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}]+\langle w_{i\tilde{j}}\rangle\langle k_{i\tilde{j}}\rangle\right)\right\}^2\\ &=\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}^2]+\langle k_{i\tilde{j}}^2\rangle-\left\{\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}]+\langle k_{i\tilde{j}}\rangle\right\}^2\\ &=v_{\mathrm{w}i}+\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}^2]-\left\{\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}]\right\}^2-2\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}]\langle k_{i\tilde{j}}\rangle\\ &=v_{\mathrm{w}i}+\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}\left[w_{ij},(k_{ij}-\langle k_{i\tilde{j}}\rangle)^2\right]-\left\{\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}]\right\}^2,\end{aligned} \tag{S33}$$

where we used the fact that $\langle k_{i\tilde{j}}^2\rangle - \langle k_{i\tilde{j}}\rangle^2 = v_{\mathrm{wi}}$. The last line of Eq. (S33) can be interpreted as the expected variance of k_{ij} within collective i after one iteration of the Wright-Fisher process excluding the effect of random sampling. Thus, let us introduce the following symbol:

$$\begin{aligned}\Delta_{\mathrm{s}} v_{\mathrm{wi}} &:= \left(\mathbb{E}_{J|I=i}[k_{IJ}^2] - \mathbb{E}_{J|I=i}[k_{IJ}]^2\right) - v_{\mathrm{wi}} \\ &= \langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}\left[w_{ij}, (k_{ij} - \langle k_{i\tilde{j}}\rangle)^2\right] - \left\{\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij}, k_{ij}]\right\}^2,\end{aligned} \tag{S34}$$

which denotes the expected change of the variance of k_{ij} within collective i due to within-collective selection.

Combining Eqs. (S28), (S32), (S33), and (S34), we can transform the first term in the last line of Eq. (S27) as follows

$$\frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[\sum_{l=1}^{n_i'}\left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right] = \frac{1}{M}\sum_{i=1}^{L}\frac{n_i\langle w_{i\tilde{j}}\rangle}{\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\left(v_{\mathrm{wi}} + \Delta_{\mathrm{s}} v_{\mathrm{wi}}\right) \tag{S35}$$

Next, we calculate the second term in the last line of Eq. (S27) as follows:

$$\begin{aligned}&\frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[n_i'\left(\langle k_{iJ_{\tilde{l}}}\rangle - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right] \\ &= \frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[n_i'\left(\frac{1}{n_i'}\sum_{l=1}^{n_i'}k_{iJ_l} - \frac{1}{n_i'}\sum_{l=1}^{n_i'}\mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right] \\ &= \frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[n_i'\left\{\frac{1}{n_i'}\sum_{l=1}^{n_i'}\left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)\right\}^2\right] \\ &= \frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[\frac{1}{n_i'}\left\{\sum_{l=1}^{n_i'}\left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)\right\}^2\right] \\ &= \frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[\frac{1}{n_i'}\left\{\sum_{l=1}^{n_i'}\left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right.\right. \\ &\qquad\left.\left. + 2\sum_{m\neq l}\left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)\left(k_{iJ_m} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)\right\}\right] \\ &= \frac{1}{M}\sum_{i=1}^{L}\mathbb{E}\left[\frac{1}{n_i'}\sum_{l=1}^{n_i'}\left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right],\end{aligned} \tag{S36}$$

where we used the fact that k_{iJ_l} and k_{iJ_m} are independent of each other for $l \neq m$ in the final step. Since n_i' can be zero, the last line of Eq. (S36) needs to interpreted as follows:

$$\frac{1}{n_i'}\sum_{l=1}^{n_i'}\left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2 = \begin{cases}0, & \text{if } n_i' = 0 \\ \frac{1}{n_i'}\sum_{l=1}^{n_i'}\left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2, & \text{if } n_i' > 0.\end{cases} \tag{S37}$$

Thus,

$$\mathbb{E}\left[\frac{1}{n_i'}\sum_{l=1}^{n_i'}\left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right] = \begin{cases}0, & \text{if } n_i' = 0 \\ \mathbb{E}_{J|I=i}[k_{IJ}^2] - \mathbb{E}_{J|I=i}[k_{IJ}]^2, & \text{if } n_i' > 0,\end{cases} \tag{S38}$$

where the case for $n'_i > 0$ follows from a calculation similar to Eqs. (S28) and (S32).

To calculate the last line of Eq. (S36), we separate the case where $n'_i > 0\ \forall i \in \{1, 2, 3, \cdots, L\}$ and the case where $n'_i = 0$ for some $i \in \{1, 2, 3, \cdots, L\}$. Let S be a proper subset of $\{1, 2, 3, \cdots, L\}$, and $P(S)$ be the probability that $n'_i = 0\ \forall i \in S$ and $n'_i > 0\ \forall i \notin S$. The value of $P(S)$ can be estimated as follows:

$$\begin{aligned} P(S) &\leqslant \left[1 - \sum_{i\in S} P(I = i)\right]^M \\ &= \left[1 - \sum_{i\in S} \frac{n_i \langle w_{i\tilde{j}}\rangle}{M\langle\langle w_{\tilde{i}\tilde{j}}\rangle\rangle}\right]^M \\ &\approx \exp\left(-\sum_{i\in S} \frac{n_i \langle w_{i\tilde{j}}\rangle}{\langle\langle w_{\tilde{i}\tilde{j}}\rangle\rangle}\right), \end{aligned} \tag{S39}$$

where the RHS of the first inequality is the probability that $n'_i = 0\ \forall i \in S$ and $n'_i \geqslant 0\ \forall i \notin S$. Using these symbols and Eq. (S34), we can express the last line of Eq. (S36) as follows:

$$\begin{aligned} &\frac{1}{M}\sum_{i=1}^{L} \mathbb{E}\left[\frac{1}{n'_i}\sum_{l=1}^{n'_i} \left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right], \\ &= \frac{1}{M}\sum_{S} P(S) \sum_{i\notin S} \left(\mathbb{E}_{J|I=i}[k_{IJ}^2] - \mathbb{E}_{J|I=i}[k_{IJ}]^2\right) \\ &= \frac{1}{M}\sum_{S} P(S) \sum_{i\notin S} (v_{\mathrm{w}i} + \Delta_{\mathrm{s}} v_{\mathrm{w}i}) \\ &= \frac{1}{M}\sum_{S} P(S) \left[\sum_{i=1}^{L} (v_{\mathrm{w}i} + \Delta_{\mathrm{s}} v_{\mathrm{w}i}) - \sum_{i\in S} (v_{\mathrm{w}i} + \Delta_{\mathrm{s}} v_{\mathrm{w}i})\right] \\ &= \frac{1}{M}\sum_{i=1}^{L} (v_{\mathrm{w}i} + \Delta_{\mathrm{s}} v_{\mathrm{w}i}) - \frac{1}{M}\sum_{S\neq\varnothing} P(S) \sum_{i\in S} (v_{\mathrm{w}i} + \Delta_{\mathrm{s}} v_{\mathrm{w}i}) \end{aligned} \tag{S40}$$

where $\sum_S$ is a summation over all possible $S \subset \{1, 2, \cdots, L\}$, and $\sum_{S\neq\varnothing}$ is the same summation excluding the case where S is empty.

We assume that the second term of the last line of Eq. (S40) is negligible for the following reasons. If $n_i \gg 1$ for some $i \in S$, then Eq. (S39) implies that $P(S) \approx 0$. Contrariwise, if the statement that $n_i \gg 1$ is false for all $i \in S$, then the value of $\sum_{i\in S}(v_{\mathrm{w}i} + \Delta_{\mathrm{s}} v_{\mathrm{w}i})$ is likely to be small because n_i is small. Under this assumption, Eq. (S40) implies that

$$\frac{1}{M}\sum_{i=1}^{L} \mathbb{E}\left[\frac{1}{n'_i}\sum_{l=1}^{n'_i} \left(k_{iJ_l} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\right] \approx \frac{1}{M}\sum_{i=1}^{L} (v_{\mathrm{w}i} + \Delta_{\mathrm{s}} v_{\mathrm{w}i}). \tag{S41}$$

The second term of Eq. (S25) is calculated as follows:

$$
\begin{aligned}
&\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n_i'}\left(\epsilon_{iJ_l}-\langle\epsilon_{iJ_{\tilde{l}}}\rangle\right)^2\right] \\
&=\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\sum_{l=1}^{n_i'}\left(\epsilon_{iJ_l}^2-2\langle\epsilon_{iJ_{\tilde{l}}}\rangle\epsilon_{iJ_l}+\langle\epsilon_{iJ_{\tilde{l}}}\rangle^2\right)\right] \\
&=\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\left(\sum_{l=1}^{n_i'}\epsilon_{iJ_l}^2-2\langle\epsilon_{iJ_{\tilde{l}}}\rangle\sum_{l=1}^{n_i'}\epsilon_{iJ_l}+\langle\epsilon_{iJ_{\tilde{l}}}\rangle^2\sum_{l=1}^{n_i'}\right)\right] \\
&=\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\left(\sum_{l=1}^{n_i'}\epsilon_{iJ_l}^2-2n_i'\langle\epsilon_{iJ_{\tilde{l}}}\rangle^2+n_i'\langle\epsilon_{iJ_{\tilde{l}}}\rangle^2\right)\right] \\
&=\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{L}\left(\sum_{l=1}^{n_i'}\epsilon_{iJ_l}^2-n_i'\langle\epsilon_{iJ_{\tilde{l}}}\rangle^2\right)\right] \\
&=\mathbb{E}\left[\frac{1}{M}\sum_{l=1}^{M}\epsilon_{I_lJ_l}^2-\frac{1}{M}\sum_{i=1}^{L}n_i'\langle\epsilon_{iJ_{\tilde{l}}}\rangle^2\right] \\
&=\mathbb{E}\left[\frac{1}{M}\sum_{l=1}^{M}\epsilon_{I_lJ_l}^2-\frac{1}{M}\sum_{i=1}^{L}n_i'\left(\frac{1}{n_i'}\sum_{l=1}^{n_i'}\epsilon_{I_lJ_l}\right)^2\right] \\
&=m\sigma-\frac{1}{M}\mathbb{E}\left[\sum_{i=1}^{L}\frac{1}{n_i'}\left(\sum_{l=1}^{n_i'}\epsilon_{I_lJ_l}\right)^2\right] \\
&=m\sigma-\frac{1}{M}\mathbb{E}\left[\sum_{i=1}^{L}\frac{1}{n_i'}\left(\sum_{l=1}^{n_i'}\epsilon_{I_lJ_l}^2+2\sum_{l\neq m}\epsilon_{I_lJ_l}\epsilon_{I_mJ_m}\right)\right] \\
&=m\sigma-\frac{1}{M}\mathbb{E}\left[\sum_{i=1}^{L}\frac{1}{n_i'}\sum_{l=1}^{n_i'}\epsilon_{I_lJ_l}^2\right] \\
&=m\sigma-\frac{1}{M}\sum_{S}P(S)\sum_{i\notin S}\mathbb{E}[\epsilon_{IJ}^2] \\
&=m\sigma-\frac{1}{M}\sum_{S}P(S)\left(\sum_{i=1}^{L}m\sigma-\sum_{i\in S}m\sigma\right) \\
&=m\sigma-\frac{L}{M}m\sigma+\frac{1}{M}\sum_{S}P(S)\sum_{i\in S}m\sigma \\
&=\left(1-\frac{L}{M}+\frac{1}{M}\sum_{S}P(S)|S|\right)m\sigma, \\
&\approx\left(1-\frac{L}{M}\right)m\sigma,
\end{aligned}
\tag{S42}
$$

where $|S|$ is the number of elements in S, and we have assumed that $\sum_S P(S)|S| \ll L$ in the last step.

Combining Eqs. (S25), (S27), (S35), (S36), (S41), and (S42), we obtain

$$
\begin{aligned}
\mathbb{E}[v'_{\mathrm{w}}] \approx\ & \frac{1}{M}\sum_{i=1}^{L}\frac{n_i\langle w_{i\tilde{j}}\rangle}{\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\{v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}\} \\
& -\frac{1}{M}\sum_{i=1}^{L}\{v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}\}+\left(1-\frac{L}{M}\right)m\sigma \\
=\ & \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}\right]+M^{-1}\sum_{i=1}^{L}n_i\left(v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}\right) \\
& -M^{-1}\sum_{i=1}^{L}n_i n_i^{-1}\{v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}\}+M^{-1}\sum_{i=1}^{L}n_i(1-n_i^{-1})m\sigma. \\
=\ & \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}\right]+M^{-1}\sum_{i=1}^{L}n_i(1-n_i^{-1})\left(v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}+m\sigma\right) \\
=\ & \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}\right]+\mathrm{ave}_{\tilde{i}}\left[(1-n_i^{-1})\left(v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}+m\sigma\right)\right] \\
=\ & \mathrm{ave}_{\tilde{i}}\left[(1-n_i^{-1})\left(v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}+m\sigma\right)\right],
\end{aligned}
\tag{S43}
$$

where we have assumed in the last step that $\langle w_{i\tilde{j}}\rangle$ and $v_{\mathrm{w}i}+\Delta_{\mathrm{s}}v_{\mathrm{w}i}$ are statistically uncorrelated as i varies, as we have assumed in Eq. (S17).

To enable the further calculation of Eq. (S43), we assume that

$$
n_i = \beta N^{-1}. \tag{S44}
$$

Under this assumption, we can transform Eq. (S43) as follows

$$
\begin{aligned}
\mathbb{E}[v'_{\mathrm{w}}] &\approx \mathrm{ave}_{\tilde{i}}\left[(1-n_i^{-1})\left(v_{\mathrm{w}i}+m\sigma+\Delta_{\mathrm{s}}v_{\mathrm{w}i}\right)\right] \\
&\approx \left(1-\beta N^{-1}\right)\mathrm{ave}_{\tilde{i}}\Big[v_{\mathrm{w}i}+m\sigma+\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}\left[w_{ij},(k_{ij}-\langle k_{i\tilde{j}}\rangle)^2\right] \\
&\qquad\qquad -\left\{\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}]\right\}^2\Big]. \\
&= \left(1-\beta N^{-1}\right)\Big(v_{\mathrm{w}}+m\sigma+\mathrm{ave}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},(k_{ij}-\langle k_{i\tilde{j}}\rangle)^2]\right]+O(s_{\mathrm{w}}^2)\Big),
\end{aligned}
\tag{S45}
$$

where we used Eq. (S16) in the last step. Expanding w_{ij} as a Taylor series around $k_{ij}=\langle k_{i\tilde{j}}\rangle$, we can show that

$$
\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},(k_{ij}-\langle k_{i\tilde{j}}\rangle)^2] = -s_{\mathrm{w}}c_{\mathrm{w}i}+O(s_{\mathrm{w}}^2), \tag{S46}
$$

where $c_{\mathrm{w}i} := \langle(k_{i\tilde{j}}-\langle k_{i\tilde{j}}\rangle)^3\rangle$, and that

$$
\mathrm{ave}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},(k_{ij}-\langle k_{i\tilde{j}}\rangle)^2]\right] = -s_{\mathrm{w}}c_{\mathrm{w}}+O(s_{\mathrm{w}}^2), \tag{S47}
$$

where $c_{\mathrm{w}} := \mathrm{ave}_{\tilde{i}}[c_{\mathrm{w}i}]$. Substituting Eq. (S47) into Eq. (S45), we obtain

$$
\mathbb{E}[v'_{\mathrm{w}}] \approx \left(1-\beta N^{-1}\right)\left(v_{\mathrm{w}}+m\sigma-s_{\mathrm{w}}c_{\mathrm{w}}+O(s_{\mathrm{w}}^2)\right). \tag{S48}
$$

3.2 Calculation of $\mathbb{E}\left[v'_{\mathrm{a}}\right]$

To calculate $\mathbb{E}\left[v'_{\mathrm{a}}\right]$, we first calculate $\mathbb{E}\left[v'_{\mathrm{t}}\right]$ and then use the fact that $\mathbb{E}\left[v'_{\mathrm{a}}\right] = \mathbb{E}\left[v'_{\mathrm{t}}\right] - \mathbb{E}\left[v'_{\mathrm{w}}\right]$.

We can express $\mathbb{E}\left[v'_{\mathrm{t}}\right]$ as follows:

$$\begin{aligned}\mathbb{E}\left[v'_{\mathrm{t}}\right] &= \mathbb{E}\left[\frac{1}{M}\sum_{l=1}^{M}\left\{k_{I_l J_l} + \epsilon_{iJ_l} - \frac{1}{M}\sum_{l=1}^{M}(k_{I_l J_l} + \epsilon_{iJ_l})\right\}^2\right] \\ &= \left(1 - \frac{1}{M}\right)\mathbb{E}\left[(k_{IJ} - \mathbb{E}[k_{IJ}])^2\right] + \left(1 - \frac{1}{M}\right) m\sigma,\end{aligned} \tag{S49}$$

where we used the fact that $k_{I_l J_l}$ and ϵ_{iJ_l} are independent of each other, as we did in Eqs. (S23), (S24), and (S25).

The first term of the last line of Eq. (S49) is calculated as follows:

$$\begin{aligned}&\mathbb{E}\left[(k_{IJ} - \mathbb{E}[k_{IJ}])^2\right] \\ &= \mathbb{E}[\Big[\left(k_{ij} - \mathbb{E}_{J|I=i}[k_{IJ}] + \mathbb{E}_{J|I=i}[k_{IJ}] - \mathbb{E}[k_{IJ}]\right)^2\Big] \\ &= \mathbb{E}\Big[\left(k_{ij} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\Big] + \mathbb{E}\Big[\left(\mathbb{E}_{J|I=i}[k_{IJ}] - \mathbb{E}[k_{IJ}]\right)^2\Big] \\ &\quad - 2\mathbb{E}\Big[\left(k_{ij} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)\left(\mathbb{E}_{J|I=i}[k_{IJ}] - \mathbb{E}[k_{IJ}]\right)\Big] \\ &= \mathbb{E}\Big[\left(k_{ij} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\Big] + \mathbb{E}\Big[\left(\mathbb{E}_{J|I=i}[k_{IJ}] - \mathbb{E}[k_{IJ}]\right)^2\Big] \\ &\quad - 2\sum_{i=1}^{L}\sum_{j=1}^{n_i} P(I=i, J=j)\left(\mathbb{E}_{J|I=i}[k_{IJ}] - \mathbb{E}[k_{IJ}]\right)\left(k_{ij} - \mathbb{E}_{J|I=i}[k_{IJ}]\right) \\ &= \mathbb{E}\Big[\left(k_{ij} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\Big] + \mathbb{E}\Big[\left(\mathbb{E}_{J|I=i}[k_{IJ}] - \mathbb{E}[k_{IJ}]\right)^2\Big] \\ &\quad - 2\sum_{i=1}^{L} P(I=i)\left(\mathbb{E}_{J|I=i}[k_{IJ}] - \mathbb{E}[k_{IJ}]\right)\sum_{j=1}^{n_i} P(J=j|I=i)\left(k_{ij} - \mathbb{E}_{J|I=i}[k_{IJ}]\right) \\ &= \mathbb{E}\Big[\left(k_{ij} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\Big] + \mathbb{E}\Big[\left(\mathbb{E}_{J|I=i}[k_{IJ}] - \mathbb{E}[k_{IJ}]\right)^2\Big]\end{aligned} \tag{S50}$$

The first term of the last line of Eq. (S50) is calculated as follows:

$$
\begin{aligned}
&\mathbb{E}\Big[\left(k_{ij} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\Big] \\
&= \mathbb{E}\Big[k_{ij}^2 - 2\mathbb{E}_{J|I=i}[k_{IJ}]k_{ij} + \mathbb{E}_{J|I=i}[k_{IJ}]^2)\Big] \\
&= \sum_{i=1}^{L}\sum_{j=1}^{n_i} P(I=i, J=j)\left(k_{ij}^2 - 2\mathbb{E}_{J|I=i}[k_{IJ}]k_{ij} + \mathbb{E}_{J|I=i}[k_{IJ}]^2\right) \\
&= \sum_{i=1}^{L} P(I=i)\Bigg(\sum_{j=1}^{n_i} P(J=j|I=i)k_{ij}^2 \\
&\quad - 2\mathbb{E}_{J|I=i}[k_{IJ}]\sum_{j=1}^{n_i} P(J=j|I=i)k_{ij} + \mathbb{E}_{J|I=i}[k_{IJ}]^2\Bigg) \\
&= \sum_{i=1}^{L} P(I=i)\left(\mathbb{E}_{J|I=i}[k_{IJ}^2] - \mathbb{E}_{J|I=i}[k_{IJ}]^2\right) \\
&= \sum_{i=1}^{L} \frac{n_i\langle w_{i\tilde{j}}\rangle}{M\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\left(\mathbb{E}_{J|I=i}[k_{IJ}^2] - \mathbb{E}_{J|I=i}[k_{IJ}]^2\right) \\
&= \sum_{i=1}^{L} \frac{n_i\langle w_{i\tilde{j}}\rangle}{M\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\left(v_{\mathrm{w}i} + \Delta_{\mathrm{s}}v_{\mathrm{w}i}\right), \\
&= \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, v_{\mathrm{w}i} + \Delta_{\mathrm{s}}v_{\mathrm{w}i}\right] + M^{-1}\sum_{i=1}^{L} n_i\left(v_{\mathrm{w}i} + \Delta_{\mathrm{s}}v_{\mathrm{w}i}\right) \\
&= \mathrm{ave}_{\tilde{i}}\left[v_{\mathrm{w}i} + \Delta_{\mathrm{s}}v_{\mathrm{w}i}\right],
\end{aligned}
\tag{S51}
$$

where we used Eq. (S34) and the assumption that $\langle w_{i\tilde{j}}\rangle$ and $v_{\mathrm{w}i} + \Delta_{\mathrm{s}}v_{\mathrm{w}i}$ are statistically uncorrelated as i varies, which has already been made in Eq. (S17). Using Eq. (S16) and (S47), we can transform the last line of Eq. (S51) as follows:

$$
\mathrm{ave}_{\tilde{i}}\left[v_{\mathrm{w}i} + \Delta_{\mathrm{s}}v_{\mathrm{w}i}\right] = v_{\mathrm{w}} - s_{\mathrm{w}}c_{\mathrm{w}} + O(s_{\mathrm{w}}^2). \tag{S52}
$$

Therefore,

$$
\mathbb{E}\Big[\left(k_{ij} - \mathbb{E}_{J|I=i}[k_{IJ}]\right)^2\Big] = v_{\mathrm{w}} - s_{\mathrm{w}}c_{\mathrm{w}} + O(s_{\mathrm{w}}^2). \tag{S53}
$$

The second term of the last line of Eq. (S50) is calculated as follows:

$$
\begin{aligned}
&\mathbb{E}\left[\left(\mathbb{E}_{J|I=i}[k_{IJ}] - \mathbb{E}[k_{IJ}]\right)^2\right] \\
&= \mathbb{E}\left[\mathbb{E}_{J|I=i}[k_{IJ}]^2 - 2\mathbb{E}[k_{IJ}]\mathbb{E}_{J|I=i}[k_{IJ}] + \mathbb{E}[k_{IJ}]^2\right] \\
&= \sum_{i=1}^{L}\sum_{j=1}^{n_i} P(I=i, J=j)\left(\mathbb{E}_{J|I=i}[k_{IJ}]^2 - 2\mathbb{E}[k_{IJ}]\mathbb{E}_{J|I=i}[k_{IJ}] + \mathbb{E}[k_{IJ}]^2\right) \\
&= \sum_{i=1}^{L} P(I=i)\mathbb{E}_{J|I=i}[k_{IJ}]^2 - 2\mathbb{E}[k_{IJ}]\sum_{i=1}^{L} P(I=i)\mathbb{E}_{J|I=i}[k_{IJ}] + \mathbb{E}[k_{IJ}]^2 \\
&= \sum_{i=1}^{L} P(I=i)\mathbb{E}_{J|I=i}[k_{IJ}]^2 - \mathbb{E}[k_{IJ}]^2
\end{aligned}
\tag{S54}
$$

The first term of the last line of Eq. (S54) is calculated as follows:

$$
\begin{aligned}
&\sum_{i=1}^{L} P(I=i)\mathbb{E}_{J|I=i}[k_{IJ}]^2 \\
&= \sum_{i=1}^{L} P(I=i)\left[\sum_{j=1}^{n_i} P(J=j|I=i)k_{ij}\right]^2 \\
&= \sum_{i=1}^{L} P(I=i)\left[\sum_{j=1}^{n_i} \frac{w_{ij}}{n_i\langle w_{i\tilde{j}}\rangle}k_{ij}\right]^2 \\
&= \sum_{i=1}^{L} P(I=i)\left[\frac{1}{\langle w_{i\tilde{j}}\rangle n_i}\sum_{j=1}^{n_i} w_{ij}k_{ij}\right]^2 \\
&= \sum_{i=1}^{L} P(I=i)\left[\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}] + \langle k_{i\tilde{j}}\rangle\right]^2 \\
&= \sum_{i=1}^{L} \frac{n_i\langle w_{i\tilde{j}}\rangle}{M\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\left[\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}] + \langle k_{i\tilde{j}}\rangle\right]^2 \\
&= \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\sum_{i=1}^{L} \frac{n_i}{M}\langle w_{i\tilde{j}}\rangle\left[\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}] + \langle k_{i\tilde{j}}\rangle\right]^2 \\
&= \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, \left\{\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}] + \langle k_{i\tilde{j}}\rangle\right\}^2\right] \\
&\quad + \mathrm{ave}_{\tilde{i}}\left[\left\{\langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}] + \langle k_{i\tilde{j}}\rangle\right\}^2\right] \\
&= \langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, \left(\langle k_{i\tilde{j}}\rangle + \Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right)^2\right] + \mathrm{ave}_{\tilde{i}}\left[\left(\langle k_{i\tilde{j}}\rangle + \Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right)^2\right],
\end{aligned}
\tag{S55}
$$

where we have used the following notation in the final step:

$$
\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle := \langle w_{i\tilde{j}}\rangle^{-1}\mathrm{cov}_{i\tilde{j}}[w_{ij},k_{ij}]. \tag{S56}
$$

The second term of the last line of Eq. (S54) is calculated as follows:

$$
\begin{aligned}
\mathbb{E}[k_{IJ}]^2 &= \left(\sum_{i=1}^{L}\sum_{j=1}^{n_i} P(I=i,J=j)k_{ij}\right)^2 \\
&= \left(\sum_{i=1}^{L} P(I=i)\sum_{j=1}^{n_i} P(J=j|I=i)k_{ij}\right)^2 \\
&= \left(\sum_{i=1}^{L} P(I=i)\mathbb{E}_{J|I=i}[k_{IJ}]\right)^2 \\
&= \left[\sum_{i=1}^{L} \frac{n_i\langle w_{i\tilde{j}}\rangle}{M\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\mathbb{E}_{J|I=i}[k_{IJ}]\right]^2.
\end{aligned}
\tag{S57}
$$

Doing the same calculation as in Eq. (S55), we can transform Eq. (S57) as follows:

$$\begin{aligned}
&\left[\sum_{i=1}^{L}\frac{n_i\langle w_{i\tilde{j}}\rangle}{M\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\mathbb{E}_{J|I=i}[k_{IJ}]\right]^2\\
&=\left[\sum_{i=1}^{L}\frac{n_i\langle w_{i\tilde{j}}\rangle}{M\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle}\left(\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right)\right]^2\\
&=\left(\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]+\mathrm{ave}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2
\end{aligned}\tag{S58}$$

Combining Eqs. (S54), (S55), and (S58), we obtain

$$\begin{aligned}
&\mathbb{E}\left[\left(\mathbb{E}_{J|I=i}[k_{IJ}]-\mathbb{E}[k_{IJ}]\right)^2\right]\\
&=\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\left(\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right)^2\right]+\mathrm{ave}_{\tilde{i}}\left[\left(\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right)^2\right]\\
&\quad-\left(\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]+\mathrm{ave}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\\
&=\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\left(\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right)^2\right]+\mathrm{ave}_{\tilde{i}}\left[\left(\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right)^2\right]\\
&\quad-\left(\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2-\left(\mathrm{ave}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\\
&\quad-2\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\mathrm{ave}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\\
&=\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\left(\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle-\mathrm{ave}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\right]\\
&\quad-\left(\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\\
&\quad+\mathrm{ave}_{\tilde{i}}\left[\left(\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right)^2\right]-\left(\mathrm{ave}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\\
&=\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\left(\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle-\mathrm{ave}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\right]\\
&\quad-\left(\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\\
&\quad+\mathrm{ave}_{\tilde{i}}\left[\left(\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle-\mathrm{ave}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\right]
\end{aligned}\tag{S59}$$

We consider each term in the last line of Eq. (S59) in terms of the order of s_{a} and s_{w}. We begin with the first term.

$$\begin{aligned}
&\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\left(\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle-\mathrm{ave}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\right]\\
&\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\left(\langle k_{i\tilde{j}}\rangle-\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle+\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle-\mathrm{ave}_{\tilde{i}}\left[\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\right]\\
&=\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\Big[\langle w_{i\tilde{j}}\rangle,\left(\langle k_{i\tilde{j}}\rangle-\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\right)^2+\left(\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle-\mathrm{ave}_{\tilde{i}}\left[\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)^2\\
&\qquad\qquad-2\left(\langle k_{i\tilde{j}}\rangle-\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\right)\left(\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle-\mathrm{ave}_{\tilde{i}}\left[\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)\Big]\\
&=\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\left(\langle k_{i\tilde{j}}\rangle-\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\right)^2\right]+O(s_{\mathrm{w}}^2)\\
&\quad-2\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle,\left(\langle k_{i\tilde{j}}\rangle-\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\right)\left(\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle-\mathrm{ave}_{\tilde{i}}\left[\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\right)\right],
\end{aligned}\tag{S60}$$

where we used Eqs. (S16) and (S17) in the last step. The last term of the last line of Eq. (S60) is zero because $\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle$ is independent of $\langle k_{i\tilde{j}}\rangle$ and $\langle w_{i\tilde{j}}\rangle$, a fact that stems from the assumptions that $\partial s_{\mathrm{w}}/\partial\langle k_{ij}\rangle = 0$ (see the main text under "Model") and that $v_{\mathrm{w}i}$ is statistically uncorrelated with $\langle w_{i\tilde{j}}\rangle$ as i varies [see Eq. (S17)]. Thus, expanding $\langle w_{i\tilde{j}}\rangle$ as a Taylor series around $\langle k_{i\tilde{j}}\rangle = \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle$, we can transform the last line of Eq. (S60) as follows:

$$\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\,\mathrm{cov}_{\tilde{i}}\Big[\langle w_{i\tilde{j}}\rangle, \big(\langle k_{i\tilde{j}}\rangle - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\big)^2\Big] + O(s_{\mathrm{w}}^2) = s_{\mathrm{a}}c_{\mathrm{a}} + O(s_{\mathrm{w}}^2) + O(s_{\mathrm{a}}^2), \tag{S61}$$

where we introduced the following symbol:

$$c_{\mathrm{a}} := \mathrm{ave}_{\tilde{i}}\left[\big(\langle k_{i\tilde{j}}\rangle - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\big)^3\right]. \tag{S62}$$

Next, we consider the second term of the last line of Eq. (S59). Equation (S15) implies that

$$\big(\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle^{-1}\mathrm{cov}_{\tilde{i}}\left[\langle w_{i\tilde{j}}\rangle, \langle k_{i\tilde{j}}\rangle + \Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\big)^2 = O(s_{\mathrm{a}}^2). \tag{S63}$$

Finally, we consider the third term of the last line of Eq. (S59) as follows:

$$\begin{aligned}
&\mathrm{ave}_{\tilde{i}}\left[\big(\langle k_{i\tilde{j}}\rangle + \Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle - \mathrm{ave}_{\tilde{i}}\left[\langle k_{i\tilde{j}}\rangle + \Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\big)^2\right]\\
&= \mathrm{ave}_{\tilde{i}}\left[\big(\langle k_{i\tilde{j}}\rangle - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle + \Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle - \mathrm{ave}_{\tilde{i}}\left[\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\big)^2\right]\\
&= \mathrm{ave}_{\tilde{i}}\Big[\,\big(\langle k_{i\tilde{j}}\rangle - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\big)^2 + \big(\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle - \mathrm{ave}_{\tilde{i}}\left[\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\big)^2\\
&\quad + 2\big(\langle k_{i\tilde{j}}\rangle - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\big)\big(\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle - \mathrm{ave}_{\tilde{i}}\left[\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\big)\,\Big]\\
&= \mathrm{ave}_{\tilde{i}}\Big[\,\big(\langle k_{i\tilde{j}}\rangle - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\big)^2\,\Big] + \mathrm{ave}_{\tilde{i}}\Big[\,\big(\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle - \mathrm{ave}_{\tilde{i}}\left[\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\big)^2\,\Big]\\
&\quad + 2\mathrm{ave}_{\tilde{i}}\Big[\,\big(\langle k_{i\tilde{j}}\rangle - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\big)\big(\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle - \mathrm{ave}_{\tilde{i}}\left[\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\big)\,\Big]\\
&= \mathrm{ave}_{\tilde{i}}\Big[\,\big(\langle k_{i\tilde{j}}\rangle - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\big)^2\,\Big] + \mathrm{ave}_{\tilde{i}}\Big[\,\big(\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle - \mathrm{ave}_{\tilde{i}}\left[\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle\right]\big)^2\,\Big]\\
&= \mathrm{ave}_{\tilde{i}}\Big[\,\big(\langle k_{i\tilde{j}}\rangle - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\big)^2\,\Big] + O(s_{\mathrm{w}}^2),\\
&= v_{\mathrm{a}} + O(s_{\mathrm{w}}^2),
\end{aligned} \tag{S64}$$

where we have assumed that $\langle k_{i\tilde{j}}\rangle$ and $\Delta_{\mathrm{s}}\langle k_{i\tilde{j}}\rangle$ are statistically uncorrelated, an assumption that is essentially the same as the assumption made in Eq. (S17) that $v_{\mathrm{w}i}$ and $\langle w_{i\tilde{j}}\rangle$ are statistically uncorrelated.

Combining Eq. (S59), (S60), (S61), (S63), and (S64), we obtain

$$\mathbb{E}\left[\big(\mathbb{E}_{J|I=i}[k_{IJ}] - \mathbb{E}[k_{IJ}]\big)^2\right] = v_{\mathrm{a}} + s_{\mathrm{a}}c_{\mathrm{a}} + O(s_{\mathrm{a}}^2) + O(s_{\mathrm{w}}^2). \tag{S65}$$

Combining Eq. (S49), (S50), (S53), and (S65), we obtain

$$\mathbb{E}\left[v_{\mathrm{t}}'\right] = \big(1 - M^{-1}\big)\big(v_{\mathrm{a}} + s_{\mathrm{a}}c_{\mathrm{a}} + v_{\mathrm{w}} - s_{\mathrm{w}}c_{\mathrm{w}} + m\sigma + O(s_{\mathrm{a}}^2) + O(s_{\mathrm{w}}^2)\big). \tag{S66}$$

Substituting Eqs. (S48) and (S66) into $\mathbb{E}\left[v_{\mathrm{a}}'\right] = \mathbb{E}\left[v_{\mathrm{t}}'\right] - \mathbb{E}\left[v_{\mathrm{w}}'\right]$, we obtain

$$\begin{aligned}
\mathbb{E}\left[v_{\mathrm{a}}'\right] &\approx \big(1 - M^{-1}\big)\big(v_{\mathrm{a}} + s_{\mathrm{a}}c_{\mathrm{a}} + v_{\mathrm{w}} - s_{\mathrm{w}}c_{\mathrm{w}} + m\sigma + O(s_{\mathrm{a}}^2) + O(s_{\mathrm{w}}^2)\big)\\
&\quad - \big(1 - \beta N^{-1}\big)\big(v_{\mathrm{w}} + m\sigma - s_{\mathrm{w}}c_{\mathrm{w}} + O(s_{\mathrm{w}}^2)\big)\\
&= \big(1 - M^{-1}\big)\big(v_{\mathrm{a}} + s_{\mathrm{a}}c_{\mathrm{a}} + O(s_{\mathrm{a}}^2) + O(s_{\mathrm{w}}^2)\big)\\
&\quad + \big(\beta N^{-1} - M^{-1}\big)\big(v_{\mathrm{w}} + m\sigma - s_{\mathrm{w}}c_{\mathrm{w}} + O(s_{\mathrm{w}}^2)\big)
\end{aligned} \tag{S67}$$

4 Estimation of $\gamma_{\rm a}$

Tsimring et al. [1] have investigated the time evolution of the probability density $p(r,t)$ of fitness r subject to mutation and selection. In this section, we show that the results of Tsimring et al. [1] imply $C \approx -0.25V^{3/2}$, where V and C are the variance and the third central moment of $p(r,t)$, respectively. This implication is consistent with our postulate $c_{\rm a} = -\gamma_{\rm a} v_{\rm a}^{3/2}$ made in Eq. (9) of the main text, where $\gamma_{\rm a}$ was measured to be about 0.25 through simulations.

Tsimring et al. [1] have considered the following equation, which describes the time evolution of $p(r,t)$:

$$\frac{\partial}{\partial t}p(r,t) = \theta(p-p_c)\left(r-\langle r\rangle\right)p(r,t) + D\frac{\partial^2}{\partial r^2}p(r,t), \tag{S68}$$

where $\theta(x)$ is the Heaviside step function, and $\langle r\rangle$ is the average fitness defined as

$$\langle f(r)\rangle = \int_{-\infty}^{\infty} f(r)p(r,t)dr,$$

and D is a diffusion constant. The first term on the RHS of Eq. (S68) describes the effect of selection; the second term, that of mutation. The Heaviside step function accounts for the fact that the probability density $p(r,t)$ must exceed a small threshold density p_c to grow because the size of a population is not infinite in reality. Tsimring et al. have shown that Eq. (S68) allows a travelling-wave solution, in which the peak of the density travels toward higher values of r, while maintaining a pulse-like shape, at a steady-state speed (denoted by v)

$$v = cD^{2/3}, \tag{S69}$$

where the value of c depends weakly on p_c and is around 4 in a wide range of p_c [1].

Multiplying both sides of Eq. (S68) with r or $(r-\langle r\rangle)^2$ and integrating over the whole range, we get

$$\begin{aligned} \frac{d}{dt}\langle r\rangle &= V - \epsilon_1, && \text{(S70)}\\ \frac{d}{dt}V &= C - \epsilon_2 + 2D, && \text{(S71)} \end{aligned}$$

where V, C, ϵ_1, and ϵ_2 are defined as follows:

$$\begin{aligned} V &= \left\langle (r-\langle r\rangle)^2\right\rangle, && \text{(S72)}\\ C &= \left\langle (r-\langle r\rangle)^3\right\rangle, && \text{(S73)}\\ \epsilon_1 &= \left\langle \theta(p_c-p)\left(r-\langle r\rangle\right)^2\right\rangle, && \text{(S74)}\\ \epsilon_2 &= \left\langle \theta(p_c-p)\left(r-\langle r\rangle\right)^3\right\rangle. && \text{(S75)} \end{aligned}$$

In obtaining Eqs. (S70) and (S71), we have assumed that the surface terms go to zero as $r \to \pm\infty$; i.e., $\lim_{r\to\pm\infty} P(r,t) = 0$, $\lim_{r\to\pm\infty} rP(r,t) = 0$, $\lim_{r\to\pm\infty} r\frac{\partial p}{\partial r} = 0$, and $\lim_{r\to\pm\infty} r^2\frac{\partial p}{\partial r} = 0$.

For a travelling-wave solution of Eq. (S68) with a constant speed v and shape, Eq. (S70) implies

$$v = V - \epsilon_1. \tag{S76}$$

Table S1: **Correspondence between Kimura's notation [2] and ours.**

Kimura's	ours	description
c	s_{a}	among-collective selection coefficient
v	m	mutation rate per generation from non-altruistic to altruistic allele
v'	m	reverse mutation rate; we assumed $v = v'$
s'	s_{w}	within-collective selection coefficient
m	0	among-collective migration rate
$2N$	$\beta^{-1}N$	number of alleles per collective; Kimura considers diploid
∞	M	total number of alleles

From Eqs. (S69) and (S76), we get

$$D = \left(\frac{V - \epsilon_1}{c}\right)^{3/2}. \tag{S77}$$

Since v and ϵ_1 are constant, Eq. (S76) implies $dV/dt = 0$. Thus, Eq. (S71) implies

$$C = -2D + \epsilon_2. \tag{S78}$$

Equations (S77) and (S78) imply

$$C = -2\left(\frac{V - \epsilon_1}{c}\right)^{3/2} + \epsilon_2 \approx -2c^{-3/2}V^{3/2},$$

where we have assumed $\epsilon_1 \ll V$ and $\epsilon_2 \ll V$ to obtain the last term. Since c is about 4 according to Tsimring et al. [1], we get

$$C \approx -0.25V^{3/2}.$$

5 Converting Kimura's notation into ours

Kimura [2] has investigated a binary-trait model of multilevel selection and shown that within-collective selection exactly balances out among-collective selection if

$$\frac{c}{v + v' + m} - 4Ns' = 0, \tag{S79}$$

where the symbols are as described in Table S1 (see also the next paragraph). Equation (S79) includes Eq. (17) of the main text as a special case. Equation (S79) appears as Eq. (27) of Ref. [2] or Eq. (4.8) of Ref. [3] as and is derived therein under the assumption that the steady-state frequency of the altruistic allele is identical to that in the absence of selection, an approximation that is expected to be valid in the limit of weak selection.

To convert Kimura's notation into ours, we assumed that the rate of mutation from a non-altruistic to an altruistic allele is identical to the rate of mutation from the altruistic to the non-altruistic allele, so that the direction of mutation is unbiased as in our quantitative-trait model. Moreover, we assumed that the migration rate among collectives is zero since our model does not consider migration. Finally, we took account of the fact that Kimura's model considers diploid as follows. In Kimura's model, each collective consists of N diploid individuals, i.e., $2N$ alleles. The number of alleles per collective can be considered as the average number of replicators per collective in our model (i.e., $\beta^{-1}N$) because Kimura's model assumes no dominance.

6 Derivation of Kimura's result through our method

In this section, we derive Eq. (17) of the main text, which gives parameter-region boundaries of the binary-trait model, using the method developed in the main text. The most important difference between the binary-trait and quantitative-trait models resides in the definition of ϵ_{IJ}. Thus, we consider only terms involving ϵ or $m\sigma$ as described below.

First, we show that the change in the definition of ϵ_{IJ} does not affect the condition for the parameter-region boundary given by Eq. (3). In the binary-trait model, $\epsilon_{IJ} = 0$ with probability $1 - m$ and $\epsilon_{IJ} = 1 - 2k_{IJ}$ with probability m (I and J are random variables taking the indices of a sampled replicator, as defined in Section 1). Thus,

$$\begin{aligned}\mathbb{E}\left[\epsilon_{IJ}\right] &= \sum_{i=1}^{L}\sum_{j=1}^{n_i} P(I=i, J=j) \int dP(\epsilon_{IJ})\epsilon_{IJ} \\ &= \sum_{i=1}^{L}\sum_{j=1}^{n_i} \frac{w_{ij}}{M\langle\!\langle w_{\tilde{i}\tilde{j}}\rangle\!\rangle} m(1-2k_{ij}) \\ &= m(1-2\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle) + O(s_\mathrm{w}^2 + s_\mathrm{a}^2 + ms_\mathrm{w} + ms_\mathrm{a}).\end{aligned} \tag{S80}$$

Therefore, Eq. (4) needs to be modified as follows:

$$\mathbb{E}\left[\Delta\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\right] = s_\mathrm{a}v_\mathrm{a} - s_\mathrm{w}v_\mathrm{w} + m(1-2\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle) + O(s_\mathrm{w}^2 + s_\mathrm{a}^2 + ms_\mathrm{w} + ms_\mathrm{a}). \tag{S81}$$

This equation, however, becomes almost identical to Eq. (4) if the parameters are on the parameter-region boundary, on which $\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle = 1/2$. Thus, the condition for the parameter-region boundary when s_a, s_w, and m are sufficiently small is the same as in the quantitative-trait model.

Next, we consider Eq. (5) and show that the change in the definition of ϵ_{IJ} makes a significant difference, which explains the difference between the binary- and quantitative-trait models. In Eq. (5), $m\sigma$ represents the difference between v_t and the variance of $k_{IJ} + \epsilon_{IJ}$. In the quantitative-trait model, this difference is simply the variance of ϵ_{IJ} because ϵ_{IJ} and k_{IJ} are independent of each other. In the binary-trait model, however, ϵ_{IJ} and k_{IJ} are not independent, and this fact affects Eq. (5), as follows. Under the assumption that $s_\mathrm{a} = s_\mathrm{w} = 0$, the variance of $k_{IJ} + \epsilon_{IJ}$ in the binary-trait model is

$$\begin{aligned}\mathbb{E}\left[(k_{IJ} + \epsilon_{IJ} - \mathbb{E}\left[k_{IJ} + \epsilon_{IJ}\right])^2\right] &= (1 - \mathbb{E}\left[k_{IJ} + \epsilon_{IJ}\right])^2 P\left(k_{IJ} + \epsilon_{IJ} = 1\right) \\ &\quad + (0 - \mathbb{E}\left[k_{IJ} + \epsilon_{IJ}\right])^2 P\left(k_{IJ} + \epsilon_{IJ} = 0\right),\end{aligned} \tag{S82}$$

where

$$\begin{aligned}P\left(k_{IJ} + \epsilon_{IJ} = 1\right) &= \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle (1-m) + \left(1 - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\right) m \\ P\left(k_{IJ} + \epsilon_{IJ} = 0\right) &= \left(1 - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\right)(1-m) + \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle m \\ \mathbb{E}\left[k_{IJ} + \epsilon_{IJ}\right] &= P\left(k_{IJ} + \epsilon_{IJ} = 1\right).\end{aligned} \tag{S83}$$

Thus,

$$\mathbb{E}\left[(k_{IJ} + \epsilon_{IJ} - \mathbb{E}\left[k_{IJ} + \epsilon_{IJ}\right])^2\right] = v_\mathrm{t} + m(1-m)\left(1 - 2\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\right)^2, \tag{S84}$$

where we used the fact that $v_\mathrm{t} = \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle(1 - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle)$. Therefore, the expected sample variance of the next generation is

$$\mathbb{E}\left[v'_\mathrm{t}\right] = (1 - M^{-1})\left[v_\mathrm{t} + m(1-m)(1 - 2\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle)^2\right]. \tag{S85}$$

Likewise, under the assumption that all collectives always consist of $\beta^{-1}N$ replicators, the expected sample variance within a collective of the next generation is

$$\mathbb{E}\left[v'_{\mathrm{w}i}\right] = (1 - \beta N^{-1})\left[v_{\mathrm{w}i} + m(1-m)(1 - 2\langle k_{i\tilde{j}}\rangle)^2\right], \tag{S86}$$

where the index of collectives i needs to be kept because $\langle k_{i\tilde{j}}\rangle$ depends on i. Averaging $\mathbb{E}\left[v'_{\mathrm{w}i}\right]$ over i, we obtain

$$\begin{aligned} \mathbb{E}\left[v'_\mathrm{w}\right] &\approx \mathrm{ave}_{\tilde{i}}\left[\mathbb{E}\left[v'_{\mathrm{w}i}\right]\right] \\ &= (1 - \beta N^{-1})\left\{v_\mathrm{w} + m(1-m)\left[(1 - 2\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle)^2 + 4v_\mathrm{a}\right]\right\}, \end{aligned} \tag{S87}$$

where we used the fact that $v_\mathrm{a} = \mathrm{ave}_{\tilde{i}}[\langle k_{i\tilde{j}}\rangle^2] - \langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle^2$. Since $\mathbb{E}[v'_\mathrm{a}] = \mathbb{E}[v'_\mathrm{t}] - \mathbb{E}[v'_\mathrm{w}]$, we obtain

$$\begin{aligned} \mathbb{E}\left[v'_\mathrm{a}\right] = &\left(1 - M^{-1}\right)v_\mathrm{a} + \left(\beta N^{-1} - M^{-1}\right)\left[v_\mathrm{w} + m(1-m)\left(1 - 2\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle\right)^2\right] \\ &- 4\left(1 - \beta N^{-1}\right)m(1-m)v_\mathrm{a}. \end{aligned} \tag{S88}$$

If the systems is on a parameter-region boundary, $\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle = 1/2$. Thus, setting $\langle\!\langle k_{\tilde{i}\tilde{j}}\rangle\!\rangle = 1/2$, we obtain

$$\mathbb{E}\left[v'_\mathrm{w}\right] = (1 - \beta N^{-1})\left[v_\mathrm{w} + 4m(1-m)v_\mathrm{a}\right] \tag{S89}$$

$$\mathbb{E}\left[v'_\mathrm{a}\right] = \left(1 - M^{-1}\right)v_\mathrm{a} + \left(\beta N^{-1} - M^{-1}\right)v_\mathrm{w} - 4\left(1 - \beta N^{-1}\right)m\left(1-m\right)v_\mathrm{a}. \tag{S90}$$

To apply the condition for the parameter-region boundary $v_\mathrm{w}/v_\mathrm{a} \approx s_\mathrm{a}/s_\mathrm{w}$, we need to calculate the ratio $v_\mathrm{w}/v_\mathrm{a}$. To this end, dividing Eq. (S89) by Eq. (S90) on each side, we obtain

$$\frac{\mathbb{E}\left[v'_\mathrm{w}\right]}{\mathbb{E}\left[v'_\mathrm{a}\right]} = \frac{(1 - \beta N^{-1})\left[\frac{v_\mathrm{w}}{v_\mathrm{a}} + 4m(1-m)\right]}{(1 - M^{-1}) + (\beta N^{-1} - M^{-1})\frac{v_\mathrm{w}}{v_\mathrm{a}} - 4(1 - \beta N^{-1})m(1-m)}. \tag{S91}$$

Assuming a steady state (i.e., $\mathbb{E}[v'_\mathrm{w}]/\mathbb{E}[v'_\mathrm{a}] = v_\mathrm{w}/v_\mathrm{a}$), we obtain

$$\frac{v_\mathrm{w}}{v_\mathrm{a}} = \frac{4m(1-m)(1 - \beta N^{-1})}{\beta N^{-1} - M^{-1}}. \tag{S92}$$

Using the condition $v_\mathrm{w}/v_\mathrm{a} \approx s_\mathrm{a}/s_\mathrm{w}$, we obtain

$$\frac{4m(1-m)(1 - \beta N^{-1})}{\beta N^{-1} - M^{-1}} \approx \frac{s_\mathrm{a}}{s_\mathrm{w}}. \tag{S93}$$

If $\beta N^{-1} \ll 1$, $m \ll 1$, and $M \to \infty$ (Kimura's model assumes that $M = \infty$), we obtain

$$\frac{s_\mathrm{a}}{s_\mathrm{w}} \approx 4m\beta^{-1}N, \tag{S94}$$

which is the same as Eq. (17).

7 Supplementary Figures

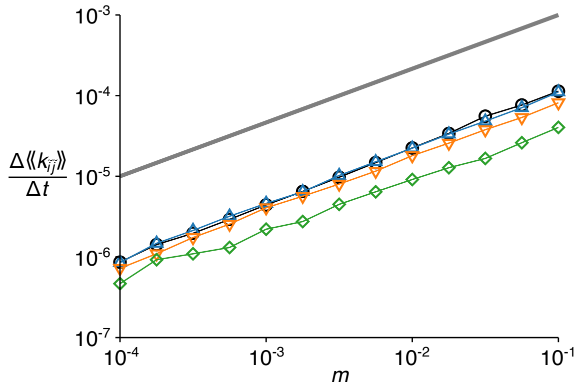

Figure S1: Rate of logarithmic fitness increase as function of mutation rate measured through simulations with no within-collective selection ($s_{\rm w} = 0$, $s_{\rm a} = 10^{-3}$, $M = 5 \times 10^5$, and $\sigma = 10^{-4}$). Fitness is defined as $w_{ij} = e^{s_{\rm a}\langle k_{i\tilde{j}}\rangle}$. Symbols have following meaning: $N = 10^2$ (black circles); $N = 10^3$ (blue triangle up); $N = 10^4$ (orange triangle down); $N = 10^5$ (green diamond). Line is $\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle/\Delta t \propto m^{2/3}$, as predicted by Eqs. (4) and (14) in main text. This figure confirms that $\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle/\Delta t \propto m^{2/3}$ in agreement with Ref. [1]. Note also that $\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle$ is roughly independent of N if $N \ll M$, which is consistent with prediction of Eq. (14) in main text.

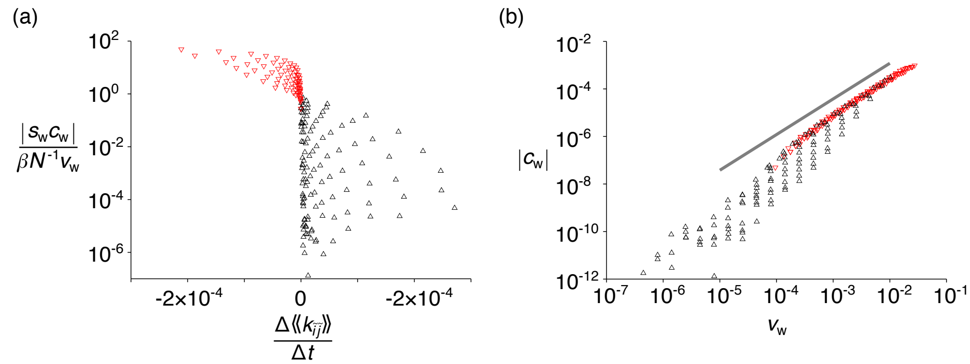

Figure S2: Average third central moment of k_{ij} within collective ($c_{\rm w}$) measured through simulations ($M = 5 \times 10^5$, $\sigma = 10^{-4}$, and $s_{\rm w} = s_{\rm a} = 10^{-2}$). (a) Ratio between effect of selection and that of random genetic drift on $\Delta v_{\rm w}$ as function of $\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle/\Delta t$: $\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle/\Delta t > 0$ (black triangle up); $\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle/\Delta t < 0$ (red triangle down). (b) $|c_{\rm w}|$ as function of $v_{\rm w}$. Triangles are simulation results: $\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle/\Delta t > 3 \times 10^{-7}$ (black triangle up); $\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle/\Delta t < 3 \times 10^{-7}$ (red triangle down). Line is $|c_{\rm w}| \propto v_{\rm w}^{3/2}$, as postulated in Eq. (9). Least squares fitting of $|c_{\rm w}| = \gamma_{\rm w} v_{\rm w}^{3/2}$ to data for $\Delta\langle\langle k_{\tilde{i}\tilde{j}}\rangle\rangle/\Delta t < 3 \times 10^{-7}$ yielded $\gamma_{\rm w} \approx 0.25$.

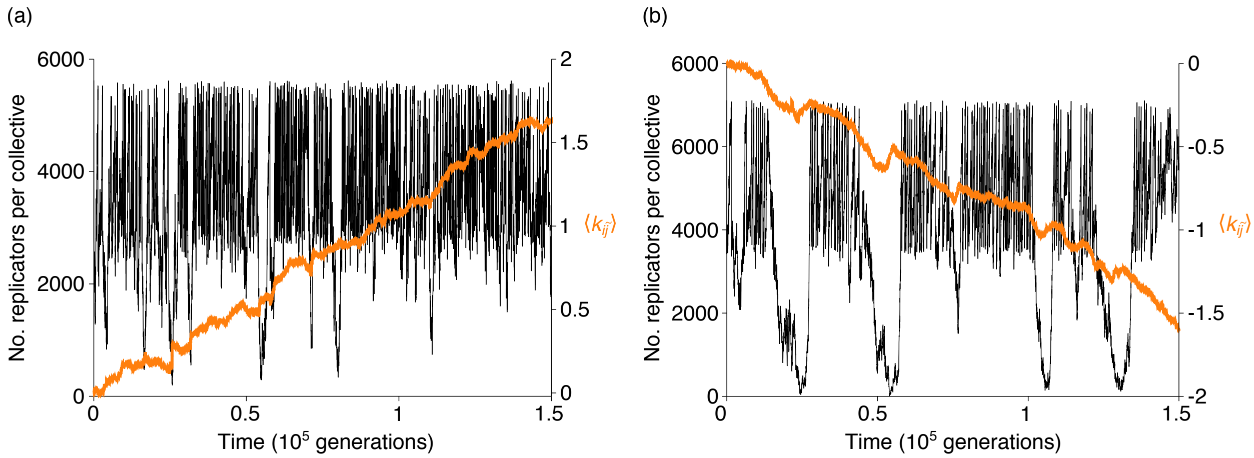

Figure S3: Dynamics of common ancestors of collectives ($m = 0.01$, $M = 5 \times 10^5$ $\sigma = 10^{-4}$, and $s_{\mathrm{a}} = s_{\mathrm{w}} = 0.01$). Plotted are number of replicators per collective (black; left coordinate) and $\langle k_{i\tilde{j}} \rangle$ (orange; right coordinate). (a) $N = 5623$. In this case, $\Delta\langle k_{\tilde{i}\tilde{j}} \rangle > 0$, and evolutionarily stable disequilibrium is not clearly observed. (b) $N = 17783$. In this case, $\Delta\langle k_{\tilde{i}\tilde{j}} \rangle < 0$, and evolutionarily stable disequilibrium is clearly observed.

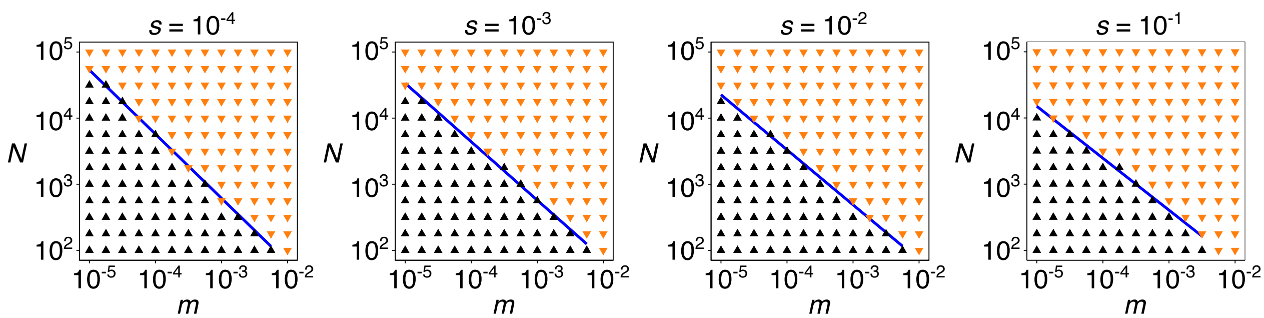

Figure S4: Parameter-sweep diagrams of binary-trait model ($s_{\mathrm{w}} = s_{\mathrm{a}} = s$ and $M = 5 \times 10^5$). Symbols have following meaning: $\langle\langle k_{\tilde{i}\tilde{j}} \rangle\rangle > 1/2$ (triangle up); $\langle\langle k_{\tilde{i}\tilde{j}} \rangle\rangle < 1/2$ (triangle down). Lines are estimated parameter-region boundaries. Parameter-region boundaries were estimated as follows. Zeros of $\langle\langle k_{\tilde{i}\tilde{j}} \rangle\rangle - 1/2$ were estimated with linear interpolation with respect to N from two simulation points around parameter-region boundary for various m values between 10^{-5} and 10^{-1}. Estimated zeros were used to obtain parameter-region boundary through least squares regression of $N \propto m^{-\alpha}$.

References

[1] Tsimring LS, Levine H, Kessler DA. RNA virus evolution via a fitness-space model. Physical Review Letters. 1996;76(23):4440-3.

[2] Kimura M. Diffusion model of population genetics incorporating group selection, with special reference to an altruistic trait. In: Ito K, Hida T, editors. Stochastic Processes and their Applications. Berlin: Springer; 1986. p. 101-18.

[3] Kimura M. Evolution of an altruistic trait through group selection as studied by the diffusion equation method. Mathematical Medicine and Biology. 1984;1(1):1-15.