# Evolutionary dynamics on any population structure 

Benjamin Allen ${ }^{1,2,3}$, Gabor Lippner ${ }^{4}$, Yu-Ting Chen ${ }^{5}$, Babak Fotouhi ${ }^{1,6}$, Naghmeh Momeni ${ }^{1,7}$, Shing-Tung Yau ${ }^{3,8}$, and Martin A. Nowak ${ }^{1,8,9}$<br>${ }^{1}$ Program for Evolutionary Dynamics, Harvard University, Cambridge, MA, USA<br>${ }^{2}$ Department of Mathematics, Emmanuel College, Boston, MA, USA<br>${ }^{3}$ Center for Mathematical Sciences and Applications, Harvard University, Cambridge, MA, USA<br>${ }^{4}$ Department of Mathematics, Northeastern University, Boston, MA, USA<br>${ }^{5}$ Department of Mathematics, University of Tennessee, Knoxville, TN, USA<br>${ }^{6}$ Institute for Quantitative Social Sciences, Harvard University, Cambridge, MA, USA<br>${ }^{7}$ Department of Electrical and Computer Engineering, McGill University, Montreal, Canada<br>${ }^{8}$ Department of Mathematics, Harvard University, Cambridge, MA, USA<br>${ }^{9}$ Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

December 23, 2016


#### Abstract

Evolution occurs in populations of reproducing individuals. The structure of a biological population affects which traits evolve [1, 2]. Understanding evolutionary game dynamics in structured populations is difficult. Precise results have been absent for a long time, but have recently emerged for special structures where all individuals have the same number of neighbors [3, 4, 5, 6, 7]. But the problem of determining which trait is favored by selection in the natural case where the number of neighbors can vary, has remained open. For arbitrary selection intensity, the problem is in a computational complexity class which suggests there is no efficient algorithm [8]. Whether there exists a simple solution for weak selection was unanswered. Here we provide, surprisingly, a general formula for weak selection that applies to any graph or social network. Our method uses coalescent theory [9, 10


and relies on calculating the meeting times of random walks [11. We can now evaluate large numbers of diverse and heterogeneous population structures for their propensity to favor cooperation. We can also study how small changes in population structure - graph surgeryaffect evolutionary outcomes. We find that cooperation flourishes most in societies that are based on strong pairwise ties.

Population structure affects ecological and evolutionary dynamics [12, 13, 14, 2]. Evolutionary graph theory [1, 3, 6] provides a mathematical tool for representing population structure: vertices correspond to individuals and edges indicate interactions. Graphs can describe spatially structured populations of bacteria, plants or animals [15], tissue architecture and differentiation in multi-cellular organisms [16, or social networks [17, 18]. Individuals reproduce into neighboring vertices according to their fitness. The graph topology affects the rate of genetic change [19] and the balance of drift versus selection [1]. The well-mixed population, which is a classical scenario for mathematical studies of evolution, is given by the complete graph.

Of particular interest is the evolution of social behavior, which can be studied using evolutionary game theory [20, 21, 22]. Evolutionary game dynamics, which are tied to ecological dynamics [21], arise whenever reproductive rates are affected by interactions with others.

In evolutionary games on graphs [24, 3, 4, 5, 6, 25, 7, individuals interact with neighbors according to a game and reproduce based on payoff (Fig. 11). A central question is to determine which strategies succeed on a given graph. It has been shown [8] that there cannot be a general closed-form solution or polynomial-time algorithm for this question, unless $\mathrm{P}=\mathrm{NP}$. To make analytical progress, one can suppose that selection is weak, meaning that the game has only a small effect on reproductive success. In this case, exact results are known for regular graphs, where each individual has the same number of neighbors [3, 4, 5, 6]. Evolutionary games on heterogenous (non-regular) graphs, which are ubiquitous in nature [26], have only been investigated using computer simulations [24, 25], heuristic approximations [27], and special cases [28, 29, 25].

Here we obtain exact results for all weighted graphs (Fig. 1a). The edge weights, $w_{i j}$, determine the frequency of game interaction and the probability of replacement between vertices $i$ and $j$. Individuals are of two types, A and


Figure 1: Evolutionary games on weighted heterogeneous graphs. a, Population structure is represented by a graph with edge weights $w_{i j}$. Here vertices are sized proportionally to their weighted degree $w_{i}=\sum_{j} w_{i j} . \mathbf{b}$, Each individual $i$ interacts with each neighbor and retains the edge-weighted average payoff $f_{i}$, shown here for the payoff matrix (6). Payoff is translated into reproductive rate $F_{i}=1+\delta f_{i}$, where $\delta$ represents the strength of selection. c, For Death-Birth updating, first a random individual $i$ is selected to be replaced; then a neighbor $j$ is chosen with probability proportional to $w_{i j} F_{j}$ to reproduce into the vacancy. This update rule induces competition between two-step neighbors (black circles). d, Coalescent theory [9, 11, 10, 23] traces ancestries back in time as random walks. The coalescence time $\tau_{i j}$ is the expected meeting time of random walks from $i$ and $j$. The meeting point (yellow circle) represents the location of the most recent common ancestor. $\mathbf{e}$, A key quantity is the probability $p_{i}$ that a randomly chosen neighbor of $i$ will chose $i$ in return, with both choices made proportionally to edge weight. Cooperation thrives when these probabilities are large.
B. The game is specified by a payoff matrix

$$
\begin{gather*}
\\
\mathrm{A}  \tag{1}\\
\mathrm{~A} \\
\mathrm{~B}\left(\begin{array}{ll}
a & b \\
c & d
\end{array}\right) .
\end{gather*}
$$

At each time step, each individual interacts with all of its neighbors. The reproductive rate of individual $i$ is given by $F_{i}=1+\delta f_{i}$, where $f_{i}$ is the edgeweighted average of the payoff that $i$ receives from its neighbors (Fig. 1b). The parameter $\delta$ represents the strength of selection. Weak selection is the regime $0<\delta \ll 1$. Neutral drift, $\delta=0$, serves as a baseline.

As update rule, we first consider Death-Birth [3] (Fig. 1c). An individual is chosen uniformly at random to be replaced; then a neighbor is chosen proportionally to reproductive rate to reproduce into the vacancy. Offspring inherit the type of their parent. Death-Birth updating is a natural scenario for genetic evolution and also translates into social settings: a random individual resolves to update its strategy; subsequently it adopts one of its neighbors' strategies proportionally to their payoff.

Over time, the population will reach the state of all $A$ or all $B$. Suppose we introduce a single $A$ individual at a vertex chosen uniformly at random in a population consisting of $B$ individuals. The fixation probability, $\rho_{A}$, is the probability of reaching all $A$ from this initial condition. Likewise, $\rho_{B}$ is the probability of reaching all $B$ when starting with a single $B$ individual in a population otherwise of $A$. Selection favors $A$ over $B$ if $\rho_{A}>\rho_{B}$.

The outcome of selection depends on the spatial assortment of types, which can be studied using coalescent theory [9, 10]. Ancestral lineages are represented as random walks [11]. A step from $i$ to $j$ occurs with probability $p_{i j}=w_{i j} / w_{i}$, where $w_{i}=\sum_{k} w_{i k}$ is the weighted degree of vertex $i$. The coalescence time $\tau_{i j}$ is the expected meeting time of two independent random walks starting at vertices $i$ and $j$ (Fig. 1d). Coalescence times can be obtained exactly and efficiently as the solution of the system of linear equations

$$
\tau_{i j}= \begin{cases}1+\frac{1}{2} \sum_{k}\left(p_{i k} \tau_{k j}+p_{j k} \tau_{i k}\right) & i \neq j  \tag{2}\\ 0 & i=j\end{cases}
$$

We show in Appendix D. 5 that the coalescence time $\tau_{i j}$ equals the expected total time during which individuals $i$ and $j$ have different types. Therefore, if $T$ is the time to absorption (fixation or extinction of the invading type),
then $T-\tau_{i j}$ is the time during which $i$ and $j$ have the same type. Of particular interest is the expected coalescence time $t_{n}$ from the two ends of an $n$-step random walk, with the initial vertex chosen proportionally to weighted degree.

Our main result holds for any payoff matrix, but we first study a donation game. Cooperators pay a cost, $c$, and provide a benefit, $b$. Defectors pay no cost and provide no benefit. This leads to the payoff matrix

$$
\begin{gather*}
\\
\mathrm{C}  \tag{3}\\
\mathrm{D}\left(\begin{array}{cc}
\mathrm{C} & \mathrm{D} \\
b-c & -c \\
b & 0
\end{array}\right),
\end{gather*}
$$

For $b>c>0$, this game is a Prisoners' Dilemma. We find that cooperation is favored over defection, $\rho_{\mathrm{C}}>1 / N>\rho_{\mathrm{D}}$, for weak selection, if and only if

$$
\begin{equation*}
-c\left(T-t_{0}\right)+b\left(T-t_{1}\right)>-c\left(T-t_{2}\right)+b\left(T-t_{3}\right) . \tag{4}
\end{equation*}
$$

Intuitively, condition (4) states that a cooperator must have a higher average payoff than a random individual two steps away. These two-step neighbors compete with the cooperator for opportunities to reproduce (Fig $1 \mathrm{~b})$. The first term, $-c\left(T-t_{0}\right)$, is the cost for being a cooperator, which is paid for the entire time, $T$, because $t_{0}=0$. The second term, $b\left(T-t_{1}\right)$, is the average benefit that the cooperator gets from its one-step neighbors. For an expected time of $T-t_{1}$, a one-step neighbor is also a cooperator. The remaining terms, $-c\left(T-t_{2}\right)+b\left(T-t_{3}\right)$, describe the average payoff of an individual two steps away. That individual pays cost $c$ whenever it is a cooperator (time $T-t_{2}$ ) and receives benefit $b$ whenever its one-step neighbors - which are three-step neighbors of the first cooperator-are cooperators (time $T-t_{3}$ ).

Time $T$ cancels in (4), leaving $-c t_{2}+b\left(t_{3}-t_{1}\right)>0$. Therefore, if $t_{3}>t_{1}$ for a given graph, cooperation is favored whenever the benefit-to-cost ratio exceeds

$$
\begin{equation*}
\left(\frac{b}{c}\right)^{*}=\frac{t_{2}}{t_{3}-t_{1}} \tag{5}
\end{equation*}
$$

The critical threshold $(b / c)^{*}$ can be obtained exactly and efficiently for any graph by solving a system of linear equations for coalescence times and substituting into equation (5). Although equation (5) is exact only for weak selection, Monte Carlo simulations (Fig. 2) show that it is highly accurate for fitness costs of up to $2.5 \%$.


Figure 2: Simulations show accuracy of our results for moderate selection. Our results are exact for weak selection. To assess accuracy for nonweak selection, we performed Monte Carlo simulations with selection strength $\delta=0.025$ and $\operatorname{cost} c=1$. This corresponds to a fitness cost of $2.5 \%$, which was empirically determined to be the cost of a cooperative behavior in yeast [30]. Markers indicate population size times frequency of fixation for a particular value of $b$ on a particular graph. Dashed lines indicate $(b / c)^{*}$ as calculated from Eq. (5). All graphs have size $N=100$. Graphs are: Barabasi-Albert [31] (BA) with linking number $m=3$, small world [32] (SW) with initial connection distance $d=3$ and edge creation probability $p=$ 0.025, Klemm-Eguiluz [33] (KE) with linking number $m=5$ and deactivation parameter $\mu=0.2$, and Holme-Kim [34] (HK) with linking number $m=2$ and triad formation parameter $P=0.2$.


Figure 3: Graphs that promote or hinder cooperation. For any graph in which each individual has one partner with edge weight $w$ and all other edges have weight 1 , the critical benefit-cost ratio converges to 1 as $w \rightarrow$ $\infty$. Thus any cooperative behavior can be favored for sufficiently large $w$. Examples include a, a disordered network; b, a weighted regular graph in which each individual has one interaction of weight $w$ and $k$ of weight 1 ; here $(b / c)^{*}=(w+k)^{2} /\left(w^{2}+k\right)$ for $N \gg k$; and $\mathbf{c}$, a complete graph in which each individual has one partner with edge weight $w$ and all other weights are 1 ; here $(b / c)^{*}=(w-2+N)^{2} /\left[(w-2)^{2}-N\right]$; cooperation is only possible if $w>\sqrt{N}+2$ (dashed vertical lines). $\mathbf{d}$, A population divided into heterogeneous islands. Edge weights are 1 between vertices on the same island and $m<1$ between vertices on different islands. Cooperation is most favored when the islands have equal size and migration is rare. e, The weighted star graph and $\mathbf{f}$, the complete bipartite graph do not support cooperation. For both graphs, all walks of odd length are equivalent; thus $t_{3}=t_{1}$ and cooperation cannot be favored according to Eq. (5).

A positive value of the critical benefit-to-cost ratio means that cooperation can be favored if it is sufficiently effective. Positive values of $(b / c)^{*}$ always exceed-but can be arbitrarily close to-unity, at which point any cooperation that produces a net benefit is favored (Fig. $3 \mathrm{a}-\mathrm{c}$ ). A negative value, which arises for $t_{3}<t_{1}$, means that cooperation cannot be favored, but spiteful behaviors, $b<0, c>0$, are favored if $b / c<(b / c)^{*}$. If $t_{3}=t_{1}$, then $(b / c)^{*}$ is infinite, and neither cooperation nor spite are favored.

Which networks best facilitate evolution of cooperation? We find that cooperation thrives when there are strong pairwise ties between individuals (Fig. $3 \mathrm{a}-\mathrm{c}$ ). To quantify this property, let $p_{i}=\sum_{j} p_{i j} p_{j i}$ be the probability that a random walk from vertex $i$ returns to $i$ on its second step (Fig. 1®). In other words, $p_{i}$ is the probability that, if $i$ choses a neighbor (for an interaction), that choice is returned by the neighbor. Cooperation succeeds best if the $p_{i}$ are large. For graphs satisfying a particular locality property (see Appendix E.2), the critical benefit-to-cost ratio becomes $(b / c)^{*}=1 / \bar{p}$, where $\bar{p}$ is a weighted average of the $p_{i}$. For unweighted regular graphs of degree $k \ll N$, a two-step walk has probability $1 / k$ to return, yielding the condition [3] $b / c>k$.

As an application, consider a population divided into islands of arbitrary sizes (Fig. 3d). Pairs on the same island are joined by edges of weight 1. Pairs on different islands are joined by edges of weight $m<1$. Our analytical results suggest that for a fixed number of islands, cooperation is most favored when the islands have equal size and $m$ is small (Appendix K.1).

Our results also specify which population structures do not support cooperation. For example, on a weighted star (Fig. 3 b), random walks alternate at each step between the hub and a leaf. Any walk of odd length is equivalent to any other. Therefore $t_{3}=t_{1}$, which makes cooperation impossible. A similar argument applies to the unweighted complete bipartite graph (Fig. 3k).

In some cases, small changes in graph topology can dramatically alter the fate of cooperation (Fig. 4). Stars do not support cooperation, but joining two stars via their hubs allows cooperation for $b / c>5 / 2$. If we modify a star by linking pairs of leaves to obtain a "ceiling fan", cooperation is favored for $b / c>8$. These examples show how targeted interventions in network structure ("graph surgery") can facilitate transitions to more cooperative societies. Of interest is also a "dense cluster" of stars all connected via their hubs which form a complete graph (Fig. $4 \mathbf{f}$ ); for this structure the critical benefit-to-cost ratio of $3 / 2$ is less than the average degree, $\bar{k}=2$.

To explore the evolutionary consequences of graph topology on a large


Figure 4: Rescuing cooperation by graph surgery. a, The star does not support cooperation; the critical benefit-to-cost ratio is infinite. b, Joining two stars via their hubs gives $(b / c)^{*}=5 / 2$. c, Joining them via two leaves gives $(b / c)^{*}=3$. d, A "ceiling fan" has $(b / c)^{*}=8$. e, A "wheel" has $(b / c)^{*}=(429+90 \sqrt{5}) / 82$. f, If we start with $m$ stars and join their hubs in a complete graph, the critical benefit-cost ratio is $(3 m-1) /(2 m-2)$, which becomes $3 / 2$ for large $m$. All $(b / c)^{*}$ values reported here hold for many leaves. Results for arbitrary sizes are given in Appendix K.2.


Figure 5: Conditions for cooperation on 1.3 million random graphs. Random graph models are Erdös-Renyi (ER), Small World (SW), BarabasiAlbert [31] (BA), random recursive graph [35] (RR), Holme-Kim 34] (HK), Klemm-Eguiliz [33] (KE), shifted-linear preferential attachment 36] (Shifted), Forest Fire [37] (FF), and meta-networks [38] of BA graphs (Island BA) and ER graphs (Island ER). Population size $N$ varies from 100 to 150. Parameter values are given in Appendix M. Values of $(b / c)^{*}$ were obtained by solving Eq. (2) and substituting into Eq. (5). a, Scatter plot of $(b / c)^{*}$ versus mean degree $k$. $71 \%$ have positive $(b / c)^{*}$ and therefore admit the possibility of cooperation. All positive $(b / c)^{*}$ values are larger than $\bar{k}$. Negative $(b / c)^{*}$ values indicate that spite $(b<0, c>0)$ can be favored. $\mathbf{b}$, Scatter plot of $(b / c)^{*}$ versus $\bar{k}_{\text {nn }}$, the expected degree of a random neighbor of a randomly chosen vertex, which has been proposed as an approximation to the critical benefit-to-cost ratio on heterogeneous graphs [27]. Although this approximation is reasonable for many graphs, there is significant variation in $(b / c)^{*}$ for each value of $\bar{k}_{\mathrm{nn}}$. The success of cooperation depends on features of the graph topology beyond the summary statistics $\bar{k}$ and $\bar{k}_{\mathrm{nn}}$.


Figure 6: Conditions for cooperation on 40000 large random graphs. Size $N$ varies from 300 to $1000.10^{4}$ graphs were generated for each of four random graph models: Erdös-Renyi [39] (ER) with edge probability $0<$ $p<0.25$, Klemm-Eguiluz [33] (KE) with linking number $3 \leq m \leq 5$ and deactivation parameter $0<\mu<0.15$, Holme-Kim [34] (HK) with linking number $2 \leq m \leq 4$ and triad formation parameter $0<P<0.15$, and a meta-network [38] of shifted-linear preferential attachment networks 36] (Island BA) with $0<p_{\text {inter }}<0.25$; see Methods for details.


Figure 7: The critical benefit-cost threshold for all graphs of size four. There are six connected, unweighted graphs of size four. Of these, only the path graph has positive $(b / c)^{*}$. Two others have infinite $(b / c)^{*}$ and three have negative $(b / c)^{*}$. There are two connected, unweighted graphs of size three (not shown): the path, which has $(b / c)^{*}=\infty$, and the complete graph (or triangle), which has $(b / c)^{*}=-2$.


Figure 8: The critical benefit-cost threshold for all graphs of size five. There are 21 connected, unweighted graphs of size five. Exact values are shown for those with $(b / c)^{*}$ positive and below 30 . Of the $(b / c)^{*}$ values, seven are positive, twelve are negative, and two are infinite.


Figure 9: The critical benefit-cost threshold for all graphs of size six. There are 112 connected, unweighted graphs of size six. Of these, 43 have positive $(b / c)^{*}, 65$ have negative $(b / c)^{*}$, and four have $(b / c)^{*}=\infty$. Numerical values are shown for those with $(b / c)^{*}$ positive and below 30 . Significantly, there are graphs with the same degree sequence (for example, $3,2,2,1,1,1$ ) but different $(b / c)^{*}$. Of the 853 graphs of size seven (not shown), 400 have positive $(b / c)^{*}, 450$ have negative $(b / c)^{*}$, and three have $(b / c)^{*}=\infty$. For all graphs of size up to seven, all positive values of $(b / c)^{*}$ are greater than the mean degree $\bar{k}$.

dolphin

macaque

zebra

wild ass

| e |  | Weighted? | $N$ | $\bar{k}$ | $(b / c)^{*}$ | $\sigma$ |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Population | Network | no | 62 | 5.1 | 6.5 | 1.36 |
| Bottlenose dolphin | frequent association | yes | 16 | - | 21.0 | 1.1 |
| Rhesus macaque | grooming | yes | 23 | - | 6.1 | 1.39 |
| Grevy's zebra | group association | yes | 28 | - | 5.6 | 1.43 |
| Asiatic wild ass | group association | no | 5253 | 6.5 | 7.7 | 1.30 |
| Framingham study | family, coworkers, friends | no | no | 4039 | 43.7 | 48.5 |
| Facebook users | online friends | no | 4159 | 6.5 | 6.6 | 1.36 |
| arXiv:gr-qc authors | co-authorship |  |  |  |  |  |

Figure 10: Results for empirical networks. The benefit-cost threshold $(b / c)^{*}$, or equivalently the structure coefficient [28, 2] $\sigma$, gives the propensity of a population structure to support cooperative and/or Pareto-efficient behaviors. These values should be interpreted in terms of specific behaviors occurring in a population, and they depend on the network ontology (that is, the meaning of links). They can be used to facilitate comparisons across populations of similar species, or to predict consequences of changes in population structure. a, Unweighted social network of frequent associations in bottlenose dolphins (Tursiops spp.) [40. b, Grooming interaction network in rhesus macaques (Macaca mulatta), weighted by grooming frequency [41. c, Weighted network of group association in Grevy's zebras (Equus grevyi) [42]. Preferred associations, which are statistically more frequent than random, are given weight 1 . Other associations occurring at least once are given weight $\epsilon \ll 1$. d, Weighted network of group association in Asiatic wild asses (onagers) 42, with the same weighting scheme as for the zebra network. For both zebra and wild ass, the unweighted networks, including every association ever observed, are dense and behave like well-mixed populations. In contrast, the weighted networks, which emphasize close ties, can promote cooperation. e, Table showing data from panels a-d as well as a social network of family, self-reported friends, and coworkers as of 1971 from the Framingham Heart Study [43, 44, a Facebook ego-network [45, and the co-authorship network for the General Relativity and Quantum Cosmology category of the arXiv preprint server 46. Average degree is reported for unweighted graphs only; for weighted graphs it is unclear which notion of degree is most relevant. Note that large $(b / c)^{*}$ ratios, which correspond to $\sigma$ values very close to one, do not mean that cooperation is never favored. Rather, the implication is that the network behaves similarly to a large well-mixed population, in which cooperation is favored for any $2 \times 2$ game with $a+b>c+d$. The donation game does not satisfy this inequality, but other cooperative interactions do [47, 48].
scale, we calculated $(b / c)^{*}$ four four ensembles of graphs: (i) 1.3 million unweighted graphs of sizes $100 \leq N \leq 150$, generated by ten random graph models (Fig. 5), (ii) 40K unweighted graphs of sizes $300 \leq N \leq 1000$ generated by four random graph models (Fig. 6), (iii) every unweighted graph of size up to seven (Fig. 7-9), and (iv) seven empirical human and animal social networks (Fig. 10). In general we find that, as the average degree, $\bar{k}$, increases, cooperation becomes increasingly difficult and eventually impossible. However, there is considerable variance in $(b / c)^{*}$ for each value of $\bar{k}$. The success of cooperation is not determined by the average degree, or even the entire degree sequence (Fig. 9).

So far we have discussed the donation game (3), but our theory extends to any pairwise game interaction of the form (6). On any graph, the condition for natural selection to favor strategy $A$ over strategy $B$, in the sense that $\rho_{\mathrm{A}}>\rho_{\mathrm{B}}$ for weak selection, can be written [28] as $\sigma a+b>c+\sigma d$. Our result implies that $\sigma=\left(-t_{1}+t_{2}+t_{3}\right) /\left(t_{1}+t_{2}-t_{3}\right)$. Therefore, the key quantity $\sigma$ can be calculated for any graph by solving a system of linear equations (Fig. 11). The value of $\sigma$ quantifies the extent to which a graph supports cooperation in a social dilemma, or the Pareto-efficient equilibrium in a coordination game [28, 2].

Our model can be extended in various ways. Birth-Death updating [3] can be studied. Total instead of average payoff can be used to compute reproductive rates [25]. Different graphs can specify interaction and replacement [49, 4, 6]. Mutation can be introduced [50, 6]. For each of these variations, we obtain the exact critical benefit-to-cost ratios (and $\sigma$ values) in terms of coalescence times (Appendix J). Additionally, appendix L provides interpretations of our results in terms of direct fitness and - in the special case of the donation game (3) -inclusive fitness.

In summary, we report here the first analytic result describing how natural selection chooses between competing strategies on any graph for weak selection. Our framework applies to strategic interactions among humans in social networks, as well as ecological and evolutionary interactions among simpler organisms in any kind of spatially structured population (Fig. 10). Our results reveal which population structures promote certain behaviors, such as cooperation. We find that cooperation flourishes most in the presence of strong pairwise ties, which is an intriguing mathematical argument for the importance of stable partnerships in forming the backbone of cooperative societies.


Figure 11: Structure coefficients for 1.3 million random graphs. We computed the structure coefficient [28] $\sigma=\left[(b / c)^{*}+1\right] /\left[(b / c)^{*}-1\right]$ for the same ensemble of random graphs as in Fig. 5 of the main text. Strategy $A$ is favored over strategy $B$ under weak selection if $\sigma a+b>c+\sigma d$; see Eq. (6) of Methods. a, Scatter plot of $\sigma$ versus $(\bar{k}+1) /(\bar{k}-1)$, which is the $\sigma$-value for a regular graph of the same mean degree $\bar{k}$. $\mathbf{b}$, Scatter plot of $\sigma$ versus $\left(\bar{k}_{\mathrm{nn}}+1\right) /\left(\bar{k}_{\mathrm{nn}}-1\right)$, which is the $\sigma$-value one would expect if the condition [27] $b / c>\bar{k}_{\mathrm{nn}}$ were exact. Here, $\bar{k}_{\mathrm{nn}}$ is the expected degree of a neighbor of a randomly chosen vertex.

## Acknowledgements

Supported by a grant from the John Templeton Foundation. The Program for Evolutionary Dynamics is supported in part by a gift from B. Wu and Eric Larson. We are grateful to Siva R. Sundaresan and Daniel I. Rubenstein for providing us with data on zebra and wild ass networks.

## A Model and notation

## A. 1 Graph

Population structure is represented by a weighted, connected graph $G$ with edge weights $w_{i j}, i, j \in G . G$ is undirected: $w_{j i}=w_{i j}$ for each $i, j \in G$. Self-loops are allowed, and are represented by the weights $w_{i i}$. Self-loops represent interaction with oneself and replacement by one's own offspring.

We define the weighted degree of vertex $i$ as $w_{i}=\sum_{j \in G} w_{i j}$. The total sum of all edge weights (counting both directions for each pair of vertices) is denoted $W$ :

$$
W=\sum_{i, j \in G} w_{i j}=\sum_{i \in G} w_{i} .
$$

## A. 2 Random walks

We will make use of random walks on $G$ in both discrete and continuous time. For a random walk on $G$, steps are taken with probability proportional to edge weight; thus the probability of a step from $i$ to $j$ is $p_{i j}=w_{i j} / w_{i}$. The probability that an $n$-step walk from vertex $i$ terminates at vertex $j$ is denoted $p_{i j}^{(n)}$.

There is a unique stationary distribution $\left\{\pi_{i}\right\}_{i \in G}$ for random walks on $G$, in which vertex $i$ has stationary probability $\pi_{i}=w_{i} / W$. This means that for each $i, j \in G, \lim _{n \rightarrow \infty} p_{i j}^{(n)}=\pi_{j}$. Random walks have the reversibility property that for each $i, j \in G, \pi_{i} p_{i j}^{(n)}=\pi_{j} p_{j i}^{(n)}$.

We will use the following shorthand: for any function $g_{i}$ on $G$, we define

$$
g_{i}^{(n)}=\sum_{j \in G} p_{i j}^{(n)} g_{j}
$$

That is, $g_{i}^{(n)}$ is the expected value of $g_{j}$ where $j$ is the terminus of an $n$-step random walk from $i$. For any function $h_{i j}$ on $G \times G$, we define

$$
h^{(n)}=\sum_{i, j \in G} \pi_{i} p_{i j}^{(n)} h_{i j} .
$$

In words, $h^{(n)}$ is the expected value of $h_{i j}$ where $i$ and $j$ are the two ends of an $n$-step random walk on $G$ started from the stationary distribution.

## A. 3 Evolutionary Markov chain

For most of this Supplement we consider a continuous-time version of the Death-Birth (DB) process. The translation from discrete to continuous time does not affect fixation probabilities or other metrics for evolutionary success.

The state of the process is represented as a binary vector $\mathbf{s}=\left(s_{i}\right)_{i \in G} \in$ $\{0,1\}^{G}$, where 0 and 1 correspond to the two types and $s_{i} \in\{0,1\}$ denotes the type of vertex $i$.

Evolution is modeled as a continuous-time Markov chain $(\mathbf{S}(t))_{t \geq 0}$ on $\{0,1\}^{G}$, which we call the evolutionary Markov chain. The dynamics of the evolutionary Markov chain are described in the following subsections.

## A. 4 Payoff and reproductive rate

The edge-weighted average payoff to vertex $i$ in state $\mathbf{s}$ is denoted $f_{i}(\mathbf{s})$. We first consider the general game

$$
\begin{array}{cc}
\mathrm{A} & \mathrm{~B} \\
\mathrm{~A}\left(\begin{array}{ll}
a & b \\
\mathrm{~B} \\
c & d
\end{array}\right), \tag{6}
\end{array}
$$

Letting 1 correspond to A and 0 to B , we have

$$
\begin{equation*}
f_{i}(\mathbf{s})=a s_{i} s_{i}^{(1)}+b s_{i}\left(1-s_{i}^{(1)}\right)+c\left(1-s_{i}\right) s_{i}^{(1)}+d\left(1-s_{i}\right)\left(1-s_{i}^{(1)}\right) . \tag{7}
\end{equation*}
$$

For the donation game

$$
\begin{gather*}
\\
\mathrm{C}  \tag{8}\\
\mathrm{D}
\end{gather*}\left(\begin{array}{cc}
b-c & \mathrm{D} \\
b & 0
\end{array}\right),
$$

letting 1 correspond to C and 0 to D, Eq. (7) reduces to

$$
\begin{equation*}
f_{i}(\mathbf{s})=-c s_{i}+b \sum_{j \in G} p_{i j} s_{j} \tag{9}
\end{equation*}
$$

The reproductive rate of vertex $i$ in state $\mathbf{s}$ is $F_{i}(\mathbf{s})=1+\delta f_{i}(\mathbf{s})$, where $\delta>0$ quantifies the strength of selection.

## A. 5 Transitions

State transitions in the evolutionary Markov chain occur via replacement events, in which the occupant of one vertex is replaced by the offspring of another. We denote by $i \rightarrow j$ the event that the occupant of $j \in G$ is replaced by the offspring of $i \in G$. Replacement events occur as Poisson processes, with rates depending on the state $\mathbf{s}$. The Death-Birth process is defined using the rates

$$
\begin{equation*}
\operatorname{Rate}[i \rightarrow j](\mathbf{s})=\frac{w_{i j} F_{i}(\mathbf{s})}{\sum_{k \in G} w_{k j} F_{k}(\mathbf{s})} \tag{10}
\end{equation*}
$$

According to Eq. (10), each vertex is replaced at overall rate 1. The conditional probability that $i$ reproduces, given that vertex $j$ is replaced, is proportional to $w_{i j} F_{i}(\mathbf{s})$. In this way, Eq. 10) defines a continuous-time analogue of the DB process described in the main text.

If the replacement event $i \rightarrow j$ occurs in state $\mathbf{s}$, and $s_{j} \neq s_{i}$, then a state transition occurs and the new state $\mathbf{s}^{\prime}$ is defined by $s_{j}^{\prime}=s_{i}$ and $s_{k}^{\prime}=s_{k}$ for all $k \neq j$. (That is, vertex $j$ inherits the type of vertex $i$, and all other vertices retain their type.) If $s_{j}=s_{i}$, then no transition occurs and $\mathbf{s}^{\prime}=\mathbf{s}$.

## B Fixation probability under weak selection

The evolutionary Markov chain has two absorbing states: the state $\mathbf{1}$ for which $s_{i}=1$ for all $i \in G$, and the state $\mathbf{0}$ for which $s_{i}=0$ for all $i \in G$. These states correspond to the fixation of types B and A, respectively. All other states of the evolutionary Markov chain are transient [51, Theorem 2]. Thus from any given initial state, the evolutionary Markov chain will eventually become absorbed in one of these two fixation states. We denote by $\rho_{\mathbf{s}_{0}}$ the fixation probability of type A from state $\mathbf{s}_{0} \in\{0,1\}^{G}$-that is,
the probability that, from initial state $\mathbf{s}_{0}$, the evolutionary Markov chain becomes absorbed in state 1.

We are interested in the behavior of these fixation probabilities under weak selection-that is, to first order in $\delta$ as $\delta \rightarrow 0^{+}$. Chen [5] derived a weakselection perturbation formula for $\rho_{\mathbf{s}_{0}}$, which applies to the birth-death and death-birth processes considered here. To provide intuition for this formula, we give a heuristic derivation here, referring to [5] for mathematical details.

Our analysis focuses on the degree-weighted frequency of type A:

$$
\hat{s}=\sum_{i \in G} \pi_{i} s_{i} .
$$

The degree weighted frequency at time $t$ is represented by the random variable

$$
\hat{S}(t)=\sum_{i \in G} \pi_{i} S_{i}(t)
$$

The main idea is that $\hat{S}(t)$ is a Martingale for the neutral process $(\delta=0)$; and that weak selection can be understood as a perturbation of this Martingale. The weighting $\pi_{i}$ of vertex $i$ can be understood as its reproductive value [52, 53]. Similar arguments have been used in other contexts [54, 55, 56, 57, 58, 59 .

Consider the evolutionary Markov chain with arbitrary initial state $\mathbf{S}(0)=$ $\mathbf{s}_{0} \in\{0,1\}^{G}$. By the Fundamental Theorem of Calculus, the expected degreeweighted frequency $\mathbb{E}_{\mathbf{s}_{0}}[\hat{S}(T)]$ at time $T>0$ satisfies

$$
\begin{equation*}
\mathbb{E}_{\mathbf{s}_{0}}[\hat{S}(T)]=\hat{s}_{0}+\int_{0}^{T} \frac{d}{d t} \mathbb{E}_{\mathbf{s}_{0}}[\hat{S}(t)] d t \tag{11}
\end{equation*}
$$

In the limit $T \rightarrow \infty$, the expected degree-weighted frequency of type 1 becomes equal to its fixation probability; therefore we have

$$
\begin{equation*}
\rho_{\mathrm{s}_{0}}=\hat{s}_{0}+\int_{0}^{\infty} \frac{d}{d t} \mathbb{E}_{\mathrm{s}_{0}}[\hat{S}(t)] d t \tag{12}
\end{equation*}
$$

We now define a state function $D(\mathbf{s})$ giving the expected instantaneous rate of change in the degree-weighted frequency of type A from state $\mathbf{s} . D(\mathbf{s})$ is defined by the relation

$$
\begin{equation*}
\mathbb{E}[\hat{S}(t+\epsilon)-\hat{S}(t) \mid \hat{S}(t)=\mathbf{s}]=D(\mathbf{s}) \epsilon+o(\epsilon) \quad\left(\epsilon \rightarrow 0^{+}\right) \tag{13}
\end{equation*}
$$

The exact form of $D(\mathbf{s})$ for DB updating is derived in Section C. Substituting in Eq. (12), we obtain

$$
\begin{equation*}
\rho_{\mathbf{s}_{0}}=\hat{s}_{0}+\int_{0}^{\infty} \mathbb{E}_{\mathbf{s}_{0}}[D(\mathbf{S}(t))] d t \tag{14}
\end{equation*}
$$

We now consider the case $\delta=0$, which represents neutral drift (no selection). We indicate neutral drift by the superscript ${ }^{\circ}$. We will show in Section C that, under DB updating, $D^{\circ}(\mathbf{s})=0$ for all $\mathbf{s} \in\{0,1\}^{G}$, meaning that $\overleftrightarrow{S}(t)$ is a Martingale for neutral drift. It follows from Eq. (14) that $\rho_{\mathrm{s}_{0}}^{\circ}=\hat{s}_{0}$; that is, the fixation probability of type A, under neutral drift, is equal to its initial degree-weighted frequency (i.e., its initial reproductive value [52, 53]). In particular, a neutral mutation arising at vertex $i$ has fixation probability $\pi_{i}$.

We now turn to weak selection; that is, we consider asymptotic expansions as $\delta \rightarrow 0^{+}$. Since $D^{\circ}(\mathbf{s})=0$, we have

$$
\begin{equation*}
D(\mathbf{s})=\delta D^{\prime}(\mathbf{s})+\mathcal{O}\left(\delta^{2}\right) \tag{15}
\end{equation*}
$$

This allows us to expand the integrand in Eq. (12):

$$
\begin{align*}
\mathbb{E}_{\mathbf{s}_{0}}[D(\mathbf{S}(t))] & =\sum_{\mathbf{s}} \mathbb{P}_{\mathbf{s}_{0}}[\mathbf{S}(t)=\mathbf{s}] D(\mathbf{s}) \\
& =\delta \sum_{\mathbf{s}} \mathbb{P}_{\mathbf{s}_{0}}^{\circ}[\mathbf{S}(t)=\mathbf{s}] D^{\prime}(\mathbf{s})+\mathcal{O}\left(\delta^{2}\right) \\
& =\delta \mathbb{E}_{\mathbf{s}_{0}}^{\circ}\left[D^{\prime}(\mathbf{S}(t))\right]+\mathcal{O}\left(\delta^{2}\right) \tag{16}
\end{align*}
$$

If we could freely interchange the expansion with the integral in Eq. 12), we would have

$$
\begin{equation*}
\rho_{\mathbf{s}_{0}}=\hat{s}_{0}+\delta \int_{0}^{\infty} \mathbb{E}_{\mathbf{s}_{0}}^{\circ}\left[D^{\prime}(\mathbf{S}(t))\right] d t+\mathcal{O}\left(\delta^{2}\right) \tag{17}
\end{equation*}
$$

A formal justification for this interchange and proof of Eq. (17), for a class of models that includes the DB process considered here, is given in Theorem 3.8 of Chen [5].

For convenience, we introduce the following notation: for any function of the state $g(\mathbf{s})$ and any initial state $\mathbf{s}_{0}$, we define

$$
\begin{equation*}
\langle g\rangle_{\mathbf{s}_{0}}^{\circ}=\int_{0}^{\infty} \mathbb{E}_{\mathbf{s}_{0}}^{\circ}[g(\mathbf{S}(t))] d t \tag{18}
\end{equation*}
$$

Thus Eq. (17) can be rewritten as

$$
\begin{equation*}
\rho_{\mathbf{s}_{0}}=\hat{s}_{0}+\delta\left\langle D^{\prime}\right\rangle_{\mathbf{s}_{0}}^{\circ}+\mathcal{O}\left(\delta^{2}\right) \tag{19}
\end{equation*}
$$

Of particular biological importance is case that type 1 initially occupies only a single vertex. In biological terms, a single indivdual of a new type has appeared in the population, e.g. by mutation or migration. We therefore focus in particular on initial states $\mathbf{s}_{0}$ with exactly one vertex of type 1 . Let $\mathbf{u}$ be the probability distribution over states that assigns probability $1 / N$ to all states $\mathbf{s}$ with exactly one vertex of type 1 , and probability zero to all other states. (That is, we suppose that the new type is equally likely to arise at each vertex. This is a natural assumption for DB updating, but not necessarily for other update rules [51, 58, 19].) We use the subscript $\mathbf{u}$ to denote the expected value of a quantity when the initial state of the evolutionary Markov chain is sampled from $\mathbf{u}$.

We define the overall fixation probability of type A as $\rho_{\mathrm{A}}=\rho_{\mathbf{u}}$, the probability that A becomes fixed when starting from a single vertex chosen with uniform probability. Taking the expectation of Eq. (19) with $\mathbf{s}_{0}$ sampled from $\mathbf{u}$, we have

$$
\begin{equation*}
\rho_{\mathrm{A}}=\frac{1}{N}+\delta\left\langle D^{\prime}\right\rangle_{\mathbf{u}}^{\circ}+\mathcal{O}\left(\delta^{2}\right) \tag{20}
\end{equation*}
$$

## C Instantaneous change under weak selection

Here we compute the expected instantaneous rate of degree-weighted frequency change $D(\mathbf{s})$ from a state $\mathbf{s}$, under DB updating and weak selection. Note that if the event $i \rightarrow j$ occurs (that is, if $j$ is replaced by the offspring of $i$ ), the resulting change in $\hat{s}$ is $\pi_{j}\left(s_{i}-s_{j}\right)$. Thus the expected instantaneous
rate of degree-weighted frequency change is given by

$$
\begin{aligned}
D(\mathbf{s}) & =\sum_{j \in G} \pi_{j}\left(-s_{j}+\sum_{i \in G} s_{i} \frac{w_{i j} F_{i}(\mathbf{s})}{\sum_{k \in G} w_{k j} F_{k}(\mathbf{s})}\right) \\
& =\sum_{i \in G} s_{i}\left(-\pi_{i}+\sum_{j \in G} \pi_{j} \frac{w_{i j} F_{i}(\mathbf{s})}{\sum_{k \in G} w_{k j} F_{k}(\mathbf{s})}\right) \\
& =\delta \sum_{i \in G} s_{i}\left(\pi_{i} f_{i}(\mathbf{s})-\pi_{i} \sum_{k \in G} p_{i k}^{(2)} f_{k}(\mathbf{s})\right)+\mathcal{O}\left(\delta^{2}\right) \\
& =\delta \sum_{i \in G} \pi_{i} s_{i}\left(f_{i}^{(0)}(\mathbf{s})-f_{i}^{(2)}(\mathbf{s})\right)+\mathcal{O}\left(\delta^{2}\right) .
\end{aligned}
$$

This shows that $D^{\circ}(\mathbf{s})=0$ for all states $\mathbf{s}$ (thus $\hat{S}(t)$ is a Martingale as claimed earlier), and

$$
\begin{equation*}
D^{\prime}(\mathbf{s})=\sum_{i \in G} \pi_{i} s_{i}\left(f_{i}^{(0)}(\mathbf{s})-f_{i}^{(2)}(\mathbf{s})\right) \tag{21}
\end{equation*}
$$

For the special case of the donation game (8), we have

$$
\begin{equation*}
D^{\prime}(\mathbf{s})=\sum_{i \in G} \pi_{i} s_{i}\left(-c\left(s_{i}^{(0)}-s_{i}^{(2)}\right)+b\left(s_{i}^{(1)}-s_{i}^{(3)}\right)\right) . \tag{22}
\end{equation*}
$$

Applying Eq. 20), the fixation probability of cooperation is given by

$$
\begin{equation*}
\rho_{\mathrm{C}}=\frac{1}{N}+\delta \sum_{i \in G} \pi_{i}\left(-c\left\langle s_{i}\left(s_{i}^{(0)}-s_{i}^{(2)}\right)\right\rangle_{\mathbf{u}}^{\circ}+b\left\langle s_{i}\left(s_{i}^{(1)}-s_{i}^{(3)}\right)\right\rangle_{\mathbf{u}}^{\circ}\right)+\mathcal{O}\left(\delta^{2}\right) \tag{23}
\end{equation*}
$$

In the following section we will show how the quantities in Eq. (23) can be computed using coalescence times.

## D Coalescing random walks

A coalescing random walk ( $C R W$ ) [11, 23] is a collection of random walks on $G$ that step independently until two walks meet (or "coalesce"), after which these two step together. This process models the tracing of ancestors backwards in time. We will consider both continuous-time and discretetime versions of the coalescing random walk, starting in either case with two walkers.

## D. 1 Continuous-time and discrete-time CRWs

In the continuous-time version, we consider a pair of walkers $(X(t), Y(t))_{t \geq 0}$ with arbitrary initial vertices $X(0)=i$ and $Y(0)=j$. Each steps at Poisson rate 1 , corresponding to the rate at which sites are replaced in the continuoustime Death-Birth process. $X(t)$ and $Y(t)$ step independently until their time of coalescence (first meeting), which is denoted $T_{\text {coal }}$. After this time, $X(t)$ and $Y(t)$ step together, so that $X(t)=Y(t)$ for all $t>T_{\text {coal }}$. Probabilities and expectations under the continuous-time coalescing random walk from $i$ and $j$ are denoted $\mathbb{P}_{(i, j)}^{\mathrm{CRW}}[]$ and $\mathbb{E}_{(i, j)}^{\mathrm{CRW}}[]$ respectively.

In the discrete-time version, we consider a pair of walkers $(X(t), Y(t))_{t=0}^{\infty}$, again with arbitrary initial vertices $X(0)=i$ and $Y(0)=j$. At each time step $t=0,1, \ldots$, if $X(t) \neq Y(t)$, one of the two walkers is chosen, with equal probability, to take a random walk step. If $X(t)=Y(t)$, then both make the same random walk step. The coalescence time $T_{\text {coal }}$ is defined as the first time $t$ for which $X(t)=Y(t)$. We use tildes to indicate the discrete-time process, so that probabilities and expectations under the discrete-time coalescing random walk from $i$ and $j$ are denoted $\tilde{\mathbb{P}}_{(i, j)}^{\mathrm{CRW}}[]$ and $\tilde{\mathbb{E}}_{(i, j)}^{\mathrm{CRW}}[]$ respectively.

We note that there is a difference in time-scales for the discrete-time and continuous-time CRWs. In the continuous-time CRW, two steps occur per unit time on average (since each walker steps at rate 1). In the discrete-time CRW, one step in total is taken per unit time.

## D. 2 Coalescence times

We denote the expected coalescence time from $i$ and $j$ in the discrete-time coalescing random walk by $\tau_{i j}=\tilde{\mathbb{E}}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}\right]$. Because coalescence times can be understood as hitting times to the diagonal of $G \times G$, and because expected hitting times are preserved under transitions between discrete and continuous time [60, §2.5.3], we have $\tau_{i j}=2 \mathbb{E}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}\right]$. (The factor of two is due to the difference in time-scales.)

Now suppose that $i$ and $j$ are the two ends of a random walk of length $n$ started from the stationary distribution. Taking the expectation of $\tau_{i j}$ over all such choices, we obtain the quantity $\tau^{(n)}$, which is denoted denoted $t_{n}$ in the main text:

$$
\begin{equation*}
t_{n}=\tau^{(n)}=\sum_{i, j \in G} \pi_{i} p_{i j}^{(n)} \tau_{i j} \tag{24}
\end{equation*}
$$

## D. 3 Assortment and coalescence

Coalescing random walks represent the ancestry of a set of individuals traced backwards in time, and can therefore be used to study assortment. Mathematically, coalescing random walks are dual to the neutral $(\delta=0)$ case of our model. This duality implies that, for any initial state $\mathbf{s}_{0}$ and any pair of types $x, y \in\{0,1\}$,

$$
\begin{equation*}
\mathbb{P}_{\mathbf{s}_{0}}^{\circ}\left[s_{i}(t)=x, s_{j}(t)=y\right]=\mathbb{P}_{(i, j)}^{\mathrm{CRW}}\left[\left(\mathbf{s}_{0}\right)_{X(t)}=x,\left(\mathbf{s}_{0}\right)_{Y(t)}=y\right] . \tag{25}
\end{equation*}
$$

In biological language, this relation says that the current occupants of vertices $i$ and $j$ have the same types as their corresponding ancestors in the initial state.

We first apply this result to an initial state $\mathbf{s}_{0}$ that has a single vertex $k$ of type 1 and all others of type $0:\left(\mathbf{s}_{0}\right)_{k}=1$ and $\left(\mathbf{s}_{0}\right)_{\ell}=0$ for all $\ell \neq k$. Consider the random variable $S_{i}(t) S_{j}(t)$, which equals one if vertices $i$ and $j$ both have type 1 at time $t$ and zero otherwise. Applying the duality relation (25), we find

$$
\begin{align*}
\mathbb{E}_{\mathrm{s}_{0}}^{\circ}\left[S_{i}(t) S_{j}(t)\right] & =\mathbb{P}_{(i, j)}^{\mathrm{CRW}}[X(t)=Y(t)=k] \\
& =\mathbb{P}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}<t, X(t)=k\right] . \tag{26}
\end{align*}
$$

That is, vertices $i$ and $j$ both have type 1 at time $t$ if and only if they are both descended from the original vertex $k$ of type 1 ; in particular, this requires that the lineages of these vertices have coalesced before time $t$ in the past.

We next consider the initial state sampled from probability distribution $\mathbf{u}$ (defined in Section B), which corresponds to setting a randomly chosen vertex to type 1 (with uniform probability) and all others vertices to type 0 . Applying Eq. (26), we have

$$
\begin{align*}
\mathbb{E}_{\mathbf{u}}^{\circ}\left[S_{i}(t) S_{j}(t)\right] & =\frac{1}{N} \sum_{k \in G} \mathbb{P}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}<t, X(t)=k\right] \\
& =\frac{1}{N} \mathbb{P}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}<t\right] \tag{27}
\end{align*}
$$

The second equality follows from the law of total probability, since we have summed over all possible values of $X(t)$.

We now move from a particular time $t$ to a time-average, using the operator $\left\rangle_{\mathbf{u}}^{\circ}\right.$. We calculate:

$$
\begin{align*}
\left\langle\frac{1}{N}-s_{i} s_{j}\right\rangle_{\mathbf{u}}^{\circ} & =\int_{0}^{\infty}\left(\frac{1}{N}-\mathbb{E}_{\mathbf{u}}^{\circ}\left[S_{i}(t) S_{j}(t)\right]\right) d t \\
& =\frac{1}{N} \int_{0}^{\infty}\left(1-\mathbb{P}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}<t\right]\right) d t \\
& =\frac{1}{N} \mathbb{E}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}\right] \\
& =\frac{\tau_{i j}}{2 N} \tag{28}
\end{align*}
$$

(The factor of 2 arises from the translation from continuous time to discrete time.) In particular, this entails that for $n_{1}, n_{2} \geq 0$

$$
\begin{equation*}
\left\langle\sum_{i \in G} \pi_{i} s_{i}\left(s_{i}^{\left(n_{1}\right)}-s_{i}^{\left(n_{2}\right)}\right)\right\rangle_{\mathbf{u}}^{\circ}=\frac{\tau^{\left(n_{2}\right)}-\tau^{\left(n_{1}\right)}}{2 N} \tag{29}
\end{equation*}
$$

## D. 4 Recurrence relations for coalescence times

The coalescence times $\tau_{i j}$ satisfy the recurrence relation

$$
\tau_{i j}= \begin{cases}0 & i=j  \tag{30}\\ 1+\frac{1}{2} \sum_{k \in G}\left(p_{i k} \tau_{j k}+p_{j k} \tau_{i k}\right) & i \neq j\end{cases}
$$

Eq. (30) is a system of $\binom{N}{2}$ linear equations. The connectedness of $G$ implies that this system has a unique solution. All coalescence times $\tau_{i j}$ can be therefore obtained in polynomial time (see Section M for a discussion of algorithms and their efficiency).

We will also make use of the concept of remeeting times. To define these we introduce a process called the remeeting random walk ( $R R W$ ), which consists of two random walks $(X(t), Y(t))$ both starting at the same vertex $i \in G$ : $X(0)=Y(0)=i$. These random walks step independently until their time $T_{\text {remeet }}$ of first remeeting, after which they step together. As in the CRW, we can consider this process in either continuous time (with each walk proceeding at rate 1) or discrete time (with one of the two walkers stepping at each timestep). We let $\tau_{i i}^{+}=\tilde{\mathbb{E}}_{i}^{\mathrm{RRW}}\left[T_{\text {remeet }}\right]$ denote the expected remeeting time in
the discrete-time RRW. ( $\tau_{i i}^{+}$is shortened to $\tau_{i}$ in the main text.) A simple recurrence argument shows that

$$
\begin{equation*}
\tau_{i i}^{+}=1+\sum_{j \in G} p_{i j} \tau_{i j} \tag{31}
\end{equation*}
$$

To derive a recurrence relation for the $\tau^{(n)}$, we expand according to Eq. (30):

$$
\begin{aligned}
\tau^{(n)} & =\sum_{i, j \in G} \pi_{i} p_{i j}^{(n)} \tau_{i j} \\
& =\sum_{\substack{i, j \in G \\
i \neq j}} \pi_{i} p_{i j}^{(n)}\left(1+\frac{1}{2} \sum_{k \in G}\left(p_{i k} \tau_{j k}+p_{j k} \tau_{i k}\right)\right) \\
& =\sum_{i, j \in G} \pi_{i} p_{i j}^{(n)}\left(1+\frac{1}{2} \sum_{k \in G}\left(p_{i k} \tau_{j k}+p_{j k} \tau_{i k}\right)\right)-\sum_{i \in G} \pi_{i} p_{i i}^{(n)}\left(1+\sum_{k \in G} p_{i k} \tau_{i k}\right) \\
& =\sum_{i, j \in G} \pi_{i} p_{i j}^{(n)}+\frac{1}{2} \sum_{i, j, k \in G} \pi_{j} p_{j i}^{(n)} p_{i k} \tau_{j k}+\frac{1}{2} \sum_{i, j, k \in G} \pi_{i} p_{i j}^{(n)} p_{j k} \tau_{i k}-\sum_{i \in G} \pi_{i} p_{i i}^{(n)} \tau_{i i}^{+} \\
& =\sum_{i \in G} \pi_{i}+\frac{1}{2} \sum_{j, k \in G} \pi_{j} p_{j k}^{(n+1)} \tau_{j k}+\frac{1}{2} \sum_{i, k \in G} \pi_{i} p_{i k}^{(n+1)} \tau_{i k}-\sum_{i \in G} \pi_{i} p_{i i}^{(n)} \tau_{i i}^{+} \\
& =1+\tau^{(n+1)}-\sum_{i \in G} \pi_{i} p_{i i}^{(n)} \tau_{i i}^{+} .
\end{aligned}
$$

We therefore have the recurrence relation

$$
\begin{equation*}
\tau^{(n+1)}=\tau^{(n)}+\sum_{i \in G} \pi_{i} p_{i i}^{(n)} \tau_{i i}^{+}-1 \tag{32}
\end{equation*}
$$

In particular, we have

$$
\begin{align*}
& \tau^{(0)}=0  \tag{33}\\
& \tau^{(1)}=\sum_{i \in G} \pi_{i} \tau_{i i}^{+}-1  \tag{34}\\
& \tau^{(2)}=\sum_{i \in G} \pi_{i} \tau_{i i}^{+}\left(1+p_{i i}^{(1)}\right)-2  \tag{35}\\
& \tau^{(3)}=\sum_{i \in G} \pi_{i} \tau_{i i}^{+}\left(1+p_{i i}^{(1)}+p_{i i}^{(2)}\right)-3 . \tag{36}
\end{align*}
$$

We note that if $G$ has no self-loops then $p_{i i}^{(1)}=0$ for all $i \in G$.
Letting $n \rightarrow \infty$ in Eq. (32), (using Cesàro means to get convergence in the case that random walks on $G$ are periodic) we obtain the identity

$$
\begin{equation*}
\sum_{i \in G} \pi_{i}^{2} \tau_{i i}^{+}=1 \tag{37}
\end{equation*}
$$

Eq. (37) can also be obtained by applying Kac's formula [60, Corollary 2.24] to the random walk on $G \times G$.

## D. 5 Time spent in different strategies

The probability that $i$ and $j$ employ different strategies at time $t$, under the neutral DB process with the uniform distribution $\mathbf{u}$ over initial placements of the mutant, can be calculated as

$$
\begin{aligned}
\mathbb{P}_{\mathbf{u}}^{\circ}\left[S_{i}(t) \neq S_{j}(t)\right]= & \frac{1}{N} \sum_{k \in G} \mathbb{P}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}>t \wedge(X(t)=k \vee Y(t)=k)\right] \\
= & \frac{1}{N} \sum_{k \in G} \mathbb{P}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}>t, X(t)=k\right] \\
& +\frac{1}{N} \sum_{k \in G} \mathbb{P}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}>t, Y(t)=k\right] \\
= & \frac{2}{N} \mathbb{P}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}>t\right] .
\end{aligned}
$$

The total expected time that $i$ and $j$ spend as different strategies can be calculated as

$$
\begin{aligned}
\int_{0}^{\infty} \mathbb{P}_{\mathbf{u}}^{\circ}\left[S_{i}(t) \neq S_{j}(t)\right] d t & =\frac{2}{N} \int_{0}^{\infty} \mathbb{P}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}>t\right] d t \\
& =\frac{2}{N} \mathbb{E}_{(i, j)}^{\mathrm{CRW}}\left[T_{\text {coal }}\right] \\
& =\frac{\tau_{i j}}{N}
\end{aligned}
$$

To connect to results described in the main text, we note that one unit of time in the continuous-time DB process corresponds, on expectation, to $N$ time-steps of the discrete-time DB process described in the main text. Thus $\tau_{i j}$ is equal to the time that $i$ and $j$ employ different strategies in the neutral
discrete-time DB process. Let $\tau_{\text {abs }}$ denote the expected time until absorption (mutant fixation or extinction) in the neutral discrete-time DB process. ( $\tau_{\text {abs }}$ is denoted $T$ in the main text.) Then $\tau_{\mathrm{abs}}-\tau_{i j}$ is the expected time that $i$ and $j$ employ the same strategy, from mutant appearance until absorption, in the neutral discrete-time DB process.

## E Conditions for success

## E. 1 General case

Combining Eqs. (23) and 29), we obtain the fixation probability of cooperators under weak selection:

$$
\begin{equation*}
\rho_{\mathrm{C}}=\frac{1}{N}+\frac{\delta}{2 N}\left(-c \tau^{(2)}+b\left(\tau^{(3)}-\tau^{(1)}\right)\right)+\mathcal{O}\left(\delta^{2}\right) \tag{38}
\end{equation*}
$$

Cooperation is favored under weak selection, in the sense that $\rho_{\mathrm{C}}>1 / N$ to first order in $\delta$, if and only if

$$
\begin{equation*}
-c \tau^{(2)}+b\left(\tau^{(3)}-\tau^{(1)}\right)>0 \tag{39}
\end{equation*}
$$

This is equivalent to Condition (1) and Eq. (2) of the main text. We show in Section $F$ that this is also the condition for $\rho_{\mathrm{D}}<1 / N$.

Using Eqs. (33)-(36), we can rewrite Condition (39) as

$$
\begin{equation*}
-c\left(\sum_{i \in G} \pi_{i} \tau_{i i}^{+}\left(1+p_{i i}^{(1)}\right)-2\right)+b\left(\sum_{i \in G} \pi_{i} \tau_{i i}^{+}\left(p_{i i}^{(1)}+p_{i i}^{(2)}\right)-2\right)>0 \tag{40}
\end{equation*}
$$

## E. 2 Asymptotic formula

Here we derive Eq. (3) of the main text, $(b / c)^{*}=1 / \bar{p}$, which is an asymptotic expression that applies in a particular limit. Consider a graph $G$ with no selfloops. Then $p_{i i}^{(1)}=0$ and we shorten $p_{i i}^{(2)}$ to $p_{i}$. Substituting in Condition (40), we obtain a critical benefit-cost ratio of

$$
\begin{equation*}
\left(\frac{b}{c}\right)^{*}=\frac{\sum_{i \in G} \pi_{i} \tau_{i i}^{+}-2}{\sum_{i \in G} \pi_{i} \tau_{i i}^{+} p_{i}-2} . \tag{41}
\end{equation*}
$$

From Eq. (37) we have

$$
1=\sum_{i \in G} \pi_{i}^{2} \tau_{i i}^{+}=\sum_{i \in G}\left(\frac{\pi_{i}}{p_{i}}\right) \pi_{i} \tau_{i i}^{+} p_{i} \leq \max _{i \in G}\left(\frac{\pi_{i}}{p_{i}}\right) \sum_{i \in G} \pi_{i} \tau_{i i}^{+} p_{i} .
$$

This implies that

$$
\begin{equation*}
\sum_{i \in G} \pi_{i} \tau_{i i}^{+} p_{i} \geq\left(\max _{i \in G}\left(\frac{\pi_{i}}{p_{i}}\right)\right)^{-1} \tag{42}
\end{equation*}
$$

Now consider an infinite sequence of graphs $\left\{G_{N}\right\}_{N=N_{0}}^{\infty}$ for which $\lim _{N \rightarrow \infty} \max _{i \in G_{N}}\left(\pi_{i} / p_{i}\right)=0$. (This is the "locality property" discussed in the main text.) For such a family, Eq. (42) implies that the sums in the numerator and denominator of Eq. (41) diverge to infinity, resulting in

$$
\begin{equation*}
\lim _{N \rightarrow \infty}(b / c)^{*}=\lim _{N \rightarrow \infty} 1 / \bar{p}, \tag{43}
\end{equation*}
$$

where $\bar{p}=\left(\sum_{i \in G_{N}} \pi_{i} \tau_{i i}^{+} p_{i}\right) /\left(\sum_{i \in G_{N}} \pi_{i} \tau_{i i}^{+}\right)$is a weighted average of the $p_{i}$ with weights $\pi_{i} \tau_{i i}^{+}$.

## E. 3 Weighted regular graphs

$G$ is a weighted regular graph if each vertex has the same distribution of weights to its neighbors. This means that the set $\left\{w_{i j}\right\}_{j \in G}$ of outgoing weights is the same for each $i \in G$. In this case, $\pi_{i}=1 / N$ for each $i \in G$, and Eq. (37) implies that $\tau_{i i}^{+}=N$ for each $i \in G$. The fixation probability of cooperators (38) becomes

$$
\begin{equation*}
\rho_{C}=\frac{1}{N}+\frac{\delta}{2 N}\left(-c\left(N+N p^{(1)}-2\right)+b\left(N p^{(1)}+N p^{(2)}-2\right)\right)+\mathcal{O}\left(\delta^{2}\right) \tag{44}
\end{equation*}
$$

Above, we have dropped the subscripts of $p_{i i}^{(1)}$ and $p_{i i}^{(2)}$ because they are the same at each vertex.

In the case that $G$ has no self-loops, $p^{(1)}=0$ and $p^{(2)}=1 / \kappa$, where $\kappa=\left(\sum_{j \in G} p_{i j}^{2}\right)^{-1}$ is the Simpson degree of $G$ [15, 6]. The fixation probability (44) simplifies to

$$
\rho_{C}=\frac{1}{N}+\frac{\delta}{2 N}(-c(N-2)+b(N / \kappa-2))+\mathcal{O}\left(\delta^{2}\right)
$$

and the critical benefit-cost ratio is

$$
\begin{equation*}
\left(\frac{b}{c}\right)^{*}=\frac{N-2}{N / \kappa-2} \tag{45}
\end{equation*}
$$

Eq. (45) generalizes previous results for graphs with transitive symmetry [4, 6] and for unweighted regular graphs [5]. In the unweighted case, the Simpson degree $\kappa$ is equal to the topological degree $k$.

## F Starting from a single defector

Suppose now that we start from a single defector (again placed uniformly at random). The game matrix can be written from the defector's point of view as

$$
\begin{gather*}
\mathrm{D}  \tag{46}\\
\mathrm{D}\left(\begin{array}{cc}
0 \\
\mathrm{C} & b \\
-c & b-c
\end{array}\right) .
\end{gather*}
$$

From Eq. (21) we see that $D^{\prime}(\mathbf{s})$ - the first-order term of the expected instantaneous rate of degree-weighted frequency change - is unaffected by the addition of a constant to the payoff matrix. Thus we can, without affecting weak-selection fixation probability, add $-b+c$ to each entry of the matrix (46), resulting in the payoff matrix

$$
\begin{gathered}
\mathrm{D} \\
\mathrm{D}\left(\begin{array}{c}
-b+c \\
\mathrm{C}\left(\begin{array}{c}
c \\
-b
\end{array}\right. \\
\mathrm{C}
\end{array}\right) .
\end{gathered}
$$

This is the original payoff matrix (8) with the roles of C and D reversed and with $b$ and $c$ replaced by their opposites. The fixation probability $\rho_{\mathrm{D}}$ is therefore obtained by replacing $b$ and $c$ by their opposites in Eq. (38):

$$
\rho_{\mathrm{D}}=\frac{1}{N}+\frac{\delta}{2 N}\left(c \tau^{(2)}-b\left(\tau^{(3)}-\tau^{(1)}\right)\right)+\mathcal{O}\left(\delta^{2}\right)
$$

In particular, this shows that Condition (39) implies $\rho_{\mathrm{C}}>1 / N>\rho_{\mathrm{D}}$, and the opposite inequality in Condition (39) implies $\rho_{\mathrm{C}}<1 / N<\rho_{\mathrm{D}}$.

## G Arbitrary $2 \times 2$ games

We now consider arbitrary $2 \times 2$ matrix games of the form (6). The general condition for success follows from the Structure Coefficient Theorem [28], which states that the condition for A to be favored over B (in the sense $\rho_{\mathrm{A}}>\rho_{\mathrm{B}}$ under weak selection) takes the form $\sigma a+b>c+\sigma d$ for some structure coefficient $\sigma$ that is independent of the game. Because of this independence, we can obtain $\sigma$ from our analysis of the simplified Prisoner's Dilemma (8):

$$
\begin{equation*}
\sigma=\frac{(b / c)^{*}+1}{(b / c)^{*}-1} . \tag{47}
\end{equation*}
$$

Combining with Condition (39), we obtain

$$
\begin{equation*}
\sigma=\frac{-\tau^{(1)}+\tau^{(2)}+\tau^{(3)}}{\tau^{(1)}+\tau^{(2)}-\tau^{(3)}} \tag{48}
\end{equation*}
$$

as stated in the main text.
We note that the condition $\sigma a+b>c+\sigma d$ tells us that $\rho_{\mathrm{A}}>\rho_{\mathrm{B}}$, but does not, in general, tell us how these fixation probabilities compare with the neutral fixation probability $1 / N$ [61]. However, for games that satisfy $a+d=b+c$ (a property known as "equal gains from switching") $\sigma a+b>c+\sigma d$ implies $\rho_{\mathrm{A}}>1 / N>\rho_{\mathrm{B}}$.

## H Bounds on $(b / c)^{*}$ and $\sigma$

Here we prove the bounds $\left|(b / c)^{*}\right|>1$ and $\sigma>0$. These will follow from the inequalities $\tau^{(1)}+\tau^{(2)} \geq \tau^{(3)}$ and $\tau^{(2)}+\tau^{(3)} \geq \tau^{(1)}$. We begin with a lemma:

Lemma 1. Let $G$ be a connected weighted graph (undirected, possibly with self-loops). For each vertex $i$ of $G$,

$$
\begin{equation*}
1+p_{i i}^{(2)} \geq 4 \pi_{i} \tag{49}
\end{equation*}
$$

Equality occurs if and only if all edges not adjcacent to $i$ have weight zero.
Proof. Let vertex $i \in G$ be arbitrary and fixed. Our proof uses a variational approach: we hold constant the weights of the edges $w_{i j}$ adjacent to $i$, but allow the weights of edges not adjacent to $i$ to vary. We will also allow the edges not adjacent to $i$ to be directed-that is, $w_{j k}$ may differ from $w_{k j}$ for $j, k \neq i$-while maintaining that the edges adjacent to $i$ be undirected ( $w_{i j}=w_{j i}$ for all $j$ ).

In this generalized setting, we denote in-degrees and out-degrees as follows:

$$
w_{j}^{\mathrm{in}}=\sum_{k \in G} w_{k j}, \quad w_{j}^{\text {out }}=\sum_{k \in G} w_{j k} .
$$

The total edge weight of $G$ is again denoted $W$ :

$$
\begin{equation*}
W=\sum_{j, k \in G} w_{j k}=\sum_{j \in G} w_{j}^{\mathrm{in}}=\sum_{j \in G} w_{j}^{\text {out }} \tag{50}
\end{equation*}
$$

For vertex $i$ we have $w_{i}^{\text {in }}=w_{i}^{\text {out }}=w_{i}$. The $\pi_{i}$ in (49) is to be understood as $w_{i} / W$. It is useful to observe the identity

$$
\begin{equation*}
W=2 w_{i}+2 \sum_{h, k \neq i} w_{h k} . \tag{51}
\end{equation*}
$$

We define the function

$$
\begin{equation*}
F\left(\left(w_{\substack{h k\\
)_{\begin{subarray}{c}{h, k \neq i \\
h \neq k} }}}\end{subarray}}\right)=w_{i}+\sum_{j \in G} \frac{w_{i j}^{2}}{w_{j}^{\text {out }}}-4 \frac{w_{i}^{2}}{W} .\right. \tag{52}
\end{equation*}
$$

Note that if all edges not adjacent to $i$ have weight zero (i.e. $w_{h k}=0$ for all $h, k \neq i$ ), then $w_{j}^{\text {out }}=w_{j i}=w_{i j}$ for each $j \neq i$ and $w_{i}=W / 2$, whereupon substituting in (52) gives $F=0$. We will prove that this is the minimum value of $F$. Dividing by $w_{i}$ will then yield

$$
1+\sum_{j \in G} \frac{w_{i j}^{2}}{w_{i} w_{j}^{\text {out }}}-4 \frac{w_{i}}{W} \geq 0
$$

which is equivalent to the desired inequality (49).
To prove that zero is the minimum value of $F$, we first observe from (51) that if $\sum_{h, k \neq i} w_{h k} \geq w_{i}$ then $w_{i} \leq W / 4$. In this case, writing the first and third terms of $F$ as $w_{i}\left(1-4 w_{i} / W\right)$, we see that $F>0$. Therefore, to minimize $F$, it suffices to restrict the domain of $F$ to the simplex

$$
\Delta=\left\{\left(w_{h k}\right)_{h, k \neq i} \mid w_{h k} \geq 0, \sum_{h, k \neq i} w_{h k}<w_{i}\right\}
$$

We now look for critical points of $F$ on the interior of $\Delta$, as well as critical points of $F$ restricted to the one of the faces of $\Delta$ where one or more of the $w_{h k}$ are zero. Such critical points of $F$ have the property that for all $h, k \neq i$, either $w_{h k}=0$ or $\frac{\partial F}{\partial w_{h k}}=0$. In the latter case, we have

$$
\begin{equation*}
0=\frac{\partial F}{\partial w_{h k}}=-\frac{w_{i h}^{2}}{\left(w_{h}^{\text {out }}\right)^{2}}+4 \frac{w_{i}^{2}}{W^{2}}, \tag{53}
\end{equation*}
$$

which implies that

$$
\begin{equation*}
\frac{w_{i h}}{w_{h}^{\text {out }}}=2 \frac{w_{i}}{W} . \tag{54}
\end{equation*}
$$

We note in particular that Eq. (54) is independent of $k$. This implies a stronger characterization of potential minimizing points of $F$ : They must satisfy the property that for all $h \neq i$, either Eq. (54) holds or else $w_{h k}=0$ for all $k \neq i$. Equivalently, for potential minimizing points of $F$, the vertices $h$ other than $i$ can be partitioned into two disjoint subsets, labeled $V_{1}$ and $V_{2}$, such that $w_{h}^{\text {out }}=w_{i h}$ for $h \in V_{1}$ and $w_{h}^{\text {out }}=\left(W / 2 w_{i}\right) w_{i h}$ for $h \in V_{2}$.

We now apply Eq. (50):

$$
\begin{aligned}
W & =w_{i}+\sum_{h \in V_{1}} w_{h}^{\text {out }}+\sum_{h \in V_{2}} w_{h}^{\text {out }} \\
& =w_{i}+\sum_{h \in V_{1}} w_{i h}+\frac{W}{2 w_{i}} \sum_{h \in V_{2}} w_{i h} \\
& =2 w_{i}+\left(\frac{W}{2 w_{i}}-1\right) \sum_{h \in V_{2}} w_{i h} .
\end{aligned}
$$

Rearranging algebraically, we obtain

$$
\left(2 w_{i}-\sum_{h \in V_{2}} w_{i h}\right)\left(W-2 w_{i}\right)=0 .
$$

The first factor on the left-hand side is necessarily positive; therefore all potential minimizing points of $F$ satisfy $w_{i}=W / 2$. In this case, all edges not adjacent to $i$ have weight zero, which we have already shown implies that $F=0$. Thus $F \geq 0$ with equality if and only if all edges not adjacent to $i$ have weight zero. Dividing Eq. (52) by $w_{i}$ completes the proof.

Positivity of the numerator and the denominator of $\sigma$ in Eq. (48) for $N \geq 3$ follows from the above lemma:

Theorem 1. For any connected weighted graph $G$, the meeting times $\tau^{(1)}$, $\tau^{(2)}$, and $\tau^{(3)}$ satisfy

$$
\begin{aligned}
& \text { (a) } \tau^{(1)}+\tau^{(2)} \geq \tau^{(3)} \\
& \text { (b) } \tau^{(2)}+\tau^{(3)} \geq \tau^{(1)} \\
& \text { (c) } \tau^{(1)}+\tau^{(3)} \geq \tau^{(2)}+2 .
\end{aligned}
$$

In (a) and (b), equality occurs if and only if $G$ has two vertices and no self-loops. In (c), equality occurs if and only if $G$ has two vertices.

Proof. For (a) we use Eq. (34)-(36):

$$
\begin{aligned}
\tau^{(3)}-\tau^{(1)} & =\sum_{i \in G} \pi_{i} \tau_{i i}^{+}\left(p_{i i}^{(1)}+p_{i i}^{(2)}\right)-2 \\
& \leq \sum_{i \in G} \pi_{i} \tau_{i i}^{+}\left(p_{i i}^{(1)}+1\right)-2 \\
& =\tau^{(2)} .
\end{aligned}
$$

Equality occurs if and only if $p_{i i}^{(2)}=1$ for all $i$, which is only possible when $G$ has two vertices and no self-loops.

For (b) and (c), we combine Eqs. (36) and (37) with Lemma 1:

$$
\begin{align*}
\tau^{(3)} & =\sum_{i \in G} \pi_{i} \tau_{i i}^{+}\left(1+p_{i i}^{(1)}+p_{i i}^{(2)}\right)-3 \\
& \geq \sum_{i \in G} \pi_{i} \tau_{i i}^{+}\left(p_{i i}^{(1)}+4 \pi_{i}\right)-3 \\
& =\sum_{i \in G} \pi_{i} \tau_{i i}^{+} p_{i i}^{(1)}+1 . \tag{55}
\end{align*}
$$

Equality in (55) attains if and only if when $1+p_{i i}^{(2)}=4 \pi_{i}$ for all vertices $i$ of $G$. By Lemma 1, this means that each vertex $i$ has the property that all edges not adjacent to $i$ have weight 0 . The only connected graphs with this property are those of size two.

Inequality (c) is now obtained by using Eqs. (34) and (35) to rewrite the last line of (55) as $\tau^{(2)}-\tau^{(1)}+2$. For (b), we rewrite the last line of Eq. (55) as $\tau^{(1)}-\tau^{(2)}+2 \sum_{i \in G} \pi_{i} \tau_{i i}^{+} p_{i i}^{(1)}$, which is greater than or equal to $\tau^{(1)}-\tau^{(2)}$ with equality if and only if $G$ has no self-loops.

We now turn to the bounds on $(b / c)^{*}$ and $\sigma$.
Corollary 1. For any graph of size $N \geq 3$, we have the bounds $\sigma>0$ and $\left|(b / c)^{*}\right|>1$. For graphs of size $N=2, \rho_{A}=\rho_{B}=1 / 2$ regardless of the game (6); thus both $(b / c)^{*}$ and $\sigma$ are undefined.

Proof. We begin with the case $N \geq 3$, for which the inequalities in Theorem 1 are strict. Inequalities (b) and (a) assert that the numerator and denominator, respectively, of Eq. (48) for $\sigma$ are positive, and therefore $\sigma>0$. The bound $\left|(b / c)^{*}\right|>1$ can be obtained by writing

$$
\left(\frac{b}{c}\right)^{*}=\frac{\sigma+1}{\sigma-1},
$$

and noting that the right-hand side has absolute value greater than 1 for $\sigma>0$.

For the case $N=2$, we observe that, from an initial state with one A and one B , fixation is determined by the first death event. Whichever type is chosen for death, the other type becomes fixed, regardless of the game. Thus $(b / c)^{*}$ and $\sigma$ are undefined.

Note that $(b / c) *$ can come arbitrarily close to 1 (equivalently, $\sigma$ can be arbitrarily large), as in Fig. 3c of the main text. It is not immediately clear whether $(b / c) *$ can come arbitrarily close to -1 (equivalently, whether $\sigma$ can be arbitrarily close to 0 ).

## I Arbitrary mutation rates

Here we relax the assumption that offspring always inherit the type of the parent. We introduce mutation of arbitrary probability $0 \leq u \leq 1$ per reproduction. With probability $1-u$, a new offspring inherits the type of the parent. Otherwise a mutation occurs, which is equally likely to result in either type. So, for example, the offspring of a type A individual has probability $1-u$ of inheriting type A, probability $u / 2$ of mutating but remaining type A, and probability $u / 2$ of mutating to type B.

For $u>0$ there is a unique stationary distribution over states of the evolutionary process, which we call the mutation-selection stationary distribution or MSS distribution [51, 6]. We use $\mathbb{E}_{\text {MSS }}$ to denote the expectation of a quantity under the MSS distribution, and $\mathbb{E}_{\mathrm{MSS}}^{\circ}$ to denote the same expectation under neutral drift $(\delta=0)$.

Following Tarnita and Taylor [58], we say that type A is favored by selection if $\mathbb{E}_{\mathrm{MSS}}[\hat{x}]>\frac{1}{2}$; that is, if A has greater degree-weighted abundance than B in the MSS distribution. Eq. (B.3) of Tarnita and Taylor [58] implies that for weak selection and DB updating, A is favored if and only if

$$
\begin{equation*}
\mathbb{E}_{\mathrm{MSS}}^{\circ}\left[D^{\prime}\right]>0 \tag{56}
\end{equation*}
$$

Evaluating $\mathbb{E}_{\mathrm{MSS}}^{\circ}\left[D^{\prime}\right]$ requires analyzing the assortment of types under the MSS distribution. This can be accomplished using the method of identity-by-descent [62, 56, 4]. Two individuals are identical by descent (IBD) if no mutation separates them from their common ancestor. We let $q_{i j}$ denote the stationary probability that the occupants of $i$ and $j$ are IBD to each other.

These IBD probabilities can be rigorously defined using the notion of the IBD-enriched Markov chain [6]. IBD probabilities on an arbitrary weighted, connected graph $G$ can be obtained as the unique solution to the system of equations [6]:

$$
q_{i j}= \begin{cases}1 & i=j  \tag{57}\\ \frac{1-u}{2} \sum_{k \in G}\left(p_{i k} q_{j k}+p_{j k} q_{i k}\right) & i \neq j\end{cases}
$$

By generalizing Lemma 3 of Allen \& Nowak [6], one can use IBD probabilities to calculate assortment under the neutral MSS distribution:

$$
\begin{equation*}
\mathbb{E}_{\mathrm{MSS}}^{\circ}\left[s_{i} s_{j}\right]=\frac{1+q_{i j}}{4} \tag{58}
\end{equation*}
$$

For the donation game (8), combining Eqs. (22), (56), and (58), we obtain that cooperation is favored under weak selection if and only if

$$
\begin{equation*}
-c\left(1-q^{(2)}\right)+b\left(q^{(1)}-q^{(3)}\right)>0 \tag{59}
\end{equation*}
$$

Above, according to our convention, we have set $q^{(n)}=\sum_{i, j \in G} \pi_{i} p_{i j}^{(n)} q_{i j}$.
For an arbitrary game of the form (6), the Structure Coefficient Theorem [28] implies that A is favored under weak selection if and only if $\sigma a+b>c+\sigma d$ where

$$
\sigma=\frac{1+q^{(1)}-q^{(2)}-q^{(3)}}{1-q^{(1)}-q^{(2)}+q^{(3)}}
$$

This result generalizes Theorem 15 of Allen \& Nowak [6] to arbitrary weighted graphs.

To connect Condition (59) to our main result (39), we apply an established connection between coalescence times and identity-by-descent probabilities [63, 59]. Since each step in the discrete-time coalescing random walk corresponds to a replacement, and mutations occur with probability $u$ per replacement, the IBD probability $q_{i j}$ has the low-mutation expansion

$$
\begin{equation*}
q_{i j}=\tilde{\mathbb{E}}_{(i, j)}^{\mathrm{CRW}}\left[(1-u)^{T_{\text {coal }}}\right]=1-u \tau_{i j}+\mathcal{O}\left(u^{2}\right) \quad(u \rightarrow 0) \tag{60}
\end{equation*}
$$

Eq. (60) can also be obtained directly by comparing Eq. (57) for $q_{i j}$ with Eq. (30) for $\tau_{i j}$. The equivalence of Conditions (39) and (59) in the lowmutation limit $(u \rightarrow 0)$ follows directly from Eq. (60).

## J Variations on the model

Here we consider three variations on the model: using accumulated rather than averaged payoffs, using different interaction and replacement graphs, and using Birth-Death rather than Death-Birth updating. In each case we obtain the exact condition for success in terms of coalescence times. Although we do not explicitly combine these variations (e.g. Birth-Death updating with accumulated payoffs), such combinations can be analyzed using straightforward combinations of the methods described here.

## J. 1 Accumulated payoffs

Accumulated payoffs means that the payoffs to vertex $i$ from its neighbors are multiplied by the corresponding edge weights and summed, without normalization. For the simplified Prisoners' Dilemma (8), the accumulated payoff to vertex $i$ in state $\mathbf{s}$ is given by

$$
\begin{aligned}
f_{i}(\mathbf{s}) & =w_{i}\left(-c s_{i}+b s_{i}^{(1)}\right) \\
& =W \pi_{i}\left(-c s_{i}+b s_{i}^{(1)}\right) .
\end{aligned}
$$

To derive $(b / c)^{*}$ for accumulated payoffs, we calculate $D^{\prime}(\mathbf{s})$, the firstorder term in the instantaneous rate of change in $\hat{s}$ from state $\mathbf{s}$, starting
with Eq. 21):

$$
\begin{aligned}
D^{\prime}(\mathbf{s}) & =\sum_{i \in G} \pi_{i} s_{i}\left(f_{i}^{(0)}(\mathbf{s})-f_{i}^{(2)}(\mathbf{s})\right) \\
& =W \sum_{i \in G} \pi_{i} s_{i}\left[\pi_{i}\left(-c s_{i}+b s_{i}^{(1)}\right)-\sum_{j \in G} p_{i j}^{(2)} \pi_{j}\left(-c s_{j}+b s_{j}^{(1)}\right)\right] \\
& =W\left[-c\left(\sum_{i \in G} \pi_{i}^{2} s_{i}^{2}-\sum_{i, j \in G} \pi_{i} p_{i j}^{(2)} \pi_{j} s_{i} s_{j}\right)+b\left(\sum_{i \in G} \pi_{i}^{2} s_{i} s_{i}^{(1)}-\sum_{i, j \in G} \pi_{i} p_{i j}^{(2)} \pi_{j} s_{i} s_{j}^{(1)}\right)\right] \\
& =W\left[-c\left(\sum_{i \in G} \pi_{i}^{2} s_{i}^{2}-\sum_{i, j \in G} \pi_{j}^{2} p_{j i}^{(2)} s_{i} s_{j}\right)+b\left(\sum_{i \in G} \pi_{i}^{2} s_{i} s_{i}^{(1)}-\sum_{i, j \in G} \pi_{j}^{2} p_{j i}^{(2)} s_{i} s_{j}^{(1)}\right)\right] \\
& =W\left[-c \sum_{i, j \in G} \pi_{i}^{2} p_{i j}^{(2)}\left(s_{i}^{2}-s_{i} s_{j}\right)+b \sum_{i, j \in G} \pi_{i}^{2} p_{i j}^{(2)}\left(s_{i} s_{i}^{(1)}-s_{j} s_{i}^{(1)}\right)\right] \\
& =W\left[-c \sum_{i, j \in G} \pi_{i}^{2} p_{i j}^{(2)}\left(s_{i}^{2}-s_{i} s_{j}\right)+b \sum_{i, j, k \in G} \pi_{i}^{2} p_{i j}^{(2)} p_{i k}\left(s_{i} s_{k}-s_{j} s_{k}\right)\right] .
\end{aligned}
$$

Now combining with Eqs. (20) and (28), we obtain the fixation probability of cooperation:

$$
\rho_{C}=\frac{1}{N}+\frac{\delta W}{2 N}\left(-c \sum_{i, j \in G} \pi_{i}^{2} p_{i j}^{(2)} \tau_{i j}+b \sum_{i, j, k \in G} \pi_{i}^{2} p_{i j}^{(2)} p_{i k}\left(\tau_{j k}-\tau_{i k}\right)\right)+\mathcal{O}\left(\delta^{2}\right) .
$$

The critical benefit-cost ratio is therefore

$$
\binom{b}{c}^{*}=\frac{\sum_{i, j \in G} \pi_{i}^{2} p_{i j}^{(2)} \tau_{i j}}{\sum_{i, j, k \in G} \pi_{i}^{2} p_{i j}^{(2)} p_{i k}\left(\tau_{j k}-\tau_{i k}\right)} .
$$

## J. 2 Different interaction and replacement graphs

Here we consider a variation in which the edge weights for game interaction differ from those for replacement [49, 64, 4, 6]. In this case, the population structure is represented by a pair of weighted graphs $(G, I)$ which have the same vertex set $V$. We assume that the replacement graph $G$ is connected (so that the population is unitary), but the interaction graph $I$ need not be.

We define an $(n, m)$-random walk to be a random walk consisting of $n$ steps according to the weights of $G$, followed by $m$ steps according to the
weights of $I$. Let $p_{i j}^{(n, m)}$ denote the probability that an $(n, m)$-random walk starting at $i$ terminates at $j$. The payoff $f_{i}(\mathbf{s})$ to vertex $i$ in state $\mathbf{s}$ can be written analogously to Eq. (9) as

$$
f_{i}(\mathbf{s})=-c s_{i}+b \sum_{j \in V} p_{i j}^{(0,1)} s_{j} .
$$

(This expression uses averaged payoffs; the extension to accumulated payoffs is straightforward.)

The results in Section $B$ carry over verbatim, with the understanding that $\pi_{i}$ is defined using the weights for the replacement graph $G$. Following the steps in Section C, we obtain

$$
\begin{equation*}
D^{\prime}(\mathbf{s})=\sum_{i \in V} \pi_{i} s_{i}\left(-c\left(s_{i}^{(0,0)}-s_{i}^{(2,0)}\right)+b\left(s_{i}^{(0,1)}-s_{i}^{(2,1)}\right)\right) . \tag{61}
\end{equation*}
$$

Above, we have adopted the notation $s_{i}^{(n, m)}=\sum_{j \in V} p_{i j}^{(n, m)} s_{j}$. Applying Eq. (28), we find an analogue of Eq. (29):

$$
\begin{equation*}
\left\langle\sum_{i \in G} \pi_{i} s_{i}\left(s_{i}^{\left(n_{1}, m_{1}\right)}-s_{i}^{\left(n_{2}, m_{2}\right)}\right)\right\rangle_{\mathbf{u}}^{\circ}=\frac{\tau^{\left(n_{2}, m_{2}\right)}-\tau^{\left(n_{1}, m_{1}\right)}}{2 N} \tag{62}
\end{equation*}
$$

where

$$
\tau^{(n, m)}=\sum_{i, j \in V} \pi_{i} p_{i j}^{(n, m)} \tau_{i j}
$$

Substituting Eqs. (61) and (62) into Eq. (23) we obtain the fixation probability of cooperation:

$$
\rho_{\mathrm{C}}=\frac{1}{N}+\frac{\delta}{2 N}\left(-c \tau^{(2,0)}+b\left(\tau^{(2,1)}-\tau^{(0,1)}\right)\right)+\mathcal{O}\left(\delta^{2}\right)
$$

The critical benefit-cost ratio is therefore

$$
\binom{b}{c}^{*}=\frac{\tau^{(2,0)}}{\tau^{(2,1)}-\tau^{(0,1)}}
$$

## J. 3 Birth-Death updating

In Birth-Death (BD) updating [3], first an individual is chosen to reproduce, with probability proportional to its reproductive rate. The offspring then
replaces a random neighbor, chosen with probability proportional to edge weight.

We study a continuous-time analogue of the birth-death process, in which site $i$ replaces site $j$ at rate $F_{i}(\mathbf{s}) p_{i j}$ :

$$
\begin{equation*}
\operatorname{Rate}[i \rightarrow j](\mathbf{s})=F_{i}(\mathbf{s}) p_{i j} . \tag{63}
\end{equation*}
$$

As for DB updating, we suppose that a new type is equally likely to arise at each vertex. This assumption is mathematically convenient, although it is arguably more natural to suppose that mutations arise in proportion to how often a vertex is replaced [51, [58, 19].

Our methods for DB updating largely carry over to BD , with some modifications that we describe here.

## J.3.1 Fixation probability under weak selection

Instead of weighting each vertex $i$ by its degree, we weight by the inverse degree $1 / w_{i}$. The inverse degree can be understood as the reproductive value of vertex $i$ under BD updating [53], and is proportional to the fixation probability of a neutral mutation arising at $i$ [53, 19]. We therefore replace $\hat{s}$ with the quantity

$$
\tilde{s}=\frac{1}{\tilde{W}} \sum_{i \in G} \frac{s_{i}}{w_{i}}, \quad \tilde{W}=\sum_{i \in G} \frac{1}{w_{i}} .
$$

The arguments of Section B carry over using $\tilde{s}$ in place of $\hat{s}$, leading to an analogue of Eq. 19):

$$
\rho_{\mathbf{s}_{0}}=\tilde{s}_{0}+\delta\left\langle D^{\prime}\right\rangle_{\mathbf{s}_{0}}^{\circ}+\mathcal{O}\left(\delta^{2}\right) .
$$

Above, $D^{\prime}(\mathbf{s})$ is the first-order term of the expected instantaneous rate of change in $\tilde{s}$, defined by the following analogues of Eqs. (13) and (15):

$$
\begin{aligned}
\mathbb{E}[\tilde{S}(t+\epsilon)-\tilde{S}(t) \mid S(t)=\mathbf{s}] & =D(\mathbf{s}) \epsilon+o(\epsilon) & & \left(\epsilon \rightarrow 0^{+}\right) \\
D(\mathbf{s}) & =\delta D^{\prime}(\mathbf{s})+\mathcal{O}\left(\delta^{2}\right) & & (\delta \rightarrow 0)
\end{aligned}
$$

When we consider the uniform distribution $\mathbf{u}$ for initial mutant appearance, Eq. 20) for $\rho_{\mathbf{u}}$ carries over verbatim.

We calculate the instantaneous rate of change $D(\mathbf{s})$ as follows:

$$
\begin{aligned}
D(\mathbf{s}) & =\frac{1}{\tilde{W}} \sum_{i, j \in G} \operatorname{Rate}[i \rightarrow j] \frac{s_{i}-s_{j}}{w_{j}} \\
& =\frac{1}{\tilde{W}} \sum_{i, j \in G} \frac{w_{i j}}{w_{i} w_{j}} F_{i}(\mathbf{s})\left(s_{i}-s_{j}\right) \\
& =\frac{\delta}{\tilde{W}} \sum_{i, j \in G} \frac{w_{i j}}{w_{i} w_{j}} f_{i}(\mathbf{s})\left(s_{i}-s_{j}\right)+\mathcal{O}\left(\delta^{2}\right) \\
& =\frac{\delta}{\tilde{W}} \sum_{i, j \in G} \frac{w_{i j}}{w_{i} w_{j}} s_{i}\left(f_{i}(\mathbf{s})-f_{j}(\mathbf{s})\right)+\mathcal{O}\left(\delta^{2}\right) \\
& =\frac{\delta}{\tilde{W}} \sum_{i, j \in G} \frac{w_{i j}}{w_{i} w_{j}}\left(-c\left(s_{i}^{2}-s_{i} s_{j}\right)+b \sum_{k \in G}\left(p_{i k} s_{i} s_{k}-p_{j k} s_{i} s_{k}\right)\right)+\mathcal{O}\left(\delta^{2}\right) .
\end{aligned}
$$

Hence we get, after swapping $i$ and $j$ second part of the second sum,

$$
\begin{equation*}
D^{\prime}(\mathbf{s})=\frac{1}{\tilde{W}}\left(-c \sum_{i, j \in G} \frac{w_{i j}}{w_{i} w_{j}}\left(s_{i}^{2}-s_{i} s_{j}\right)+b \sum_{i, j, k \in G} \frac{w_{i j}}{w_{i} w_{j}} p_{i k}\left(s_{i} s_{k}-s_{j} s_{k}\right)\right) \tag{64}
\end{equation*}
$$

## J.3.2 Coalescence and assortment

The final key step is finding the appropriate modification of the coalescing random walk (CRW) used in Section D. The rate of stepping from $i$ to $j$ must correspond to the rate at which the $j \rightarrow i$ replacement occurs at neutrality, which is $w_{i j} / w_{j}$ for BD updating. Therefore the CRW for BD is a continuoustime process in which steps from $i$ to $j$ occur at rate $w_{i j} / w_{j}$ (instead of at the usual rate $\left.p_{i j}=w_{i j} / w_{i}\right)$.

We let $\tilde{\tau}_{i j}$ denote the coalescence time from vertices $i$ and $j$ under this modified CRW. (Note that the $\tilde{\tau}_{i j}$ are defined in continuous time) These coalescence times satisfy the system of equations

$$
\tilde{\tau}_{i j}= \begin{cases}0 & i=j  \tag{65}\\ \frac{1}{\sum_{\ell \in G} \frac{w_{i \ell}+w_{j \ell}}{w_{\ell}}}\left(1+\sum_{k \in G} \frac{w_{i k} \tilde{\tau}_{i k}+w_{j k} \tilde{\tau}_{j k}}{w_{k}}\right) & i \neq j .\end{cases}
$$

Again, the connectedness of $G$ implies that this system has a unique solution, which can be obtained in polynomial time (see Section $M$ ).

Armed with this modified CRW, the arguments of Section D. 3 can be readily adapted to BD updating. In particular, Eq. (28) has the analogue

$$
\begin{equation*}
\left\langle\frac{1}{N}-s_{i} s_{j}\right\rangle_{\mathbf{u}}^{\circ}=\frac{\tilde{\tau}_{i j}}{N} \tag{66}
\end{equation*}
$$

In contrast to Eq. (28), there is no factor of 2 in the denominator on the right-hand side of Eq. (66). This is because the $\tilde{\tau}_{i j}$ are defined in continuous time, whereas the $\tau_{i j}$ are defined in discrete time.

## J.3.3 Condition for success

Combining Eqs. (20), (64) and (66), we obtain the fixation probability of cooperation

$$
\rho_{C}=\frac{1}{N}+\frac{\delta}{\tilde{W} N}\left(-c \sum_{i, j \in G} \frac{w_{i j}}{w_{i} w_{j}} \tilde{\tau}_{i j}+b \sum_{i, j, k \in G} \frac{w_{i j}}{w_{i} w_{j}} p_{i k}\left(\tilde{\tau}_{j k}-\tilde{\tau}_{i k}\right)\right)+\mathcal{O}\left(\delta^{2}\right) .
$$

The critical benefit-cost ratio is therefore

$$
\left(\frac{b}{c}\right)^{*}=\frac{\sum_{i, j \in G} \frac{w_{i j}}{w_{i} w_{j}} \tilde{\tau}_{i j}}{\sum_{i, j, k \in G} \frac{w_{i j}}{w_{i} w_{j}} p_{i k}\left(\tilde{\tau}_{j k}-\tilde{\tau}_{i k}\right)} .
$$

## K Examples

## K. 1 Island model

Consider a population subdivided into $n$ islands of size $N_{1}, \ldots, N_{n}$. Weights are 1 for distinct vertices on the same island and $m<1$ for vertices on different islands. Therefore, the weighted degree of a vertex on island $i$ is $w_{i}=N_{i}-1+m\left(N-N_{i}\right)$.

We let $\tau_{i i^{\prime}}$ denote the coalescence time for two distinct vertices on island $i$, and $\tau_{i j}$ the coalescence time for a pair on islands $i$ and $j$. The recurrence
relations (30) become

$$
\begin{align*}
\tau_{i i^{\prime}}=1+ & \frac{1}{N_{i}-1+m\left(N-N_{i}\right)}\left(\left(N_{i}-2\right) \tau_{i i^{\prime}}+m \sum_{j \neq i} N_{j} \tau_{i j}\right) \\
\tau_{i j}=1+ & \frac{\left(N_{i}-1\right) \tau_{i j}+m \sum_{k \neq i, j} N_{k} \tau_{k j}+m\left(N_{j}-1\right) \tau_{j j^{\prime}}}{2\left(N_{i}-1+m\left(N-N_{i}\right)\right)}  \tag{67}\\
& +\frac{\left(N_{j}-1\right) \tau_{i j}+m \sum_{k \neq i, j} N_{k} \tau_{i k}+m\left(N_{i}-1\right) \tau_{i i^{\prime}}}{2\left(N_{j}-1+m\left(N-N_{j}\right)\right)}
\end{align*}
$$

## K.1.1 Two islands

In the case of two islands, we obtain the critical $b / c$ ratio exactly, by solving Eq. (67) with the aid of Mathematica and applying Eq. (41). We obtain an answer in the form $(b / c)^{*}=$ num/denom, where the numerator and denominator are given respectively by

$$
\begin{aligned}
& \quad \text { num }=\left((N m+N-2)^{2}-D^{2}(m-1)^{2}\right)^{2} \\
& \times\left\{D^{6}(m-1)^{3} m-D^{4}(m-1)^{2}\left(3 N^{2} m(m+1)-2 N\left(m^{2}+5 m-2\right)+10 m-6\right)\right. \\
& +D^{2} N(m-1)\left[3 N^{3} m(m+1)^{2}-2 N^{2}\left(2 m^{3}+9 m^{2}+6 m-1\right)\right. \\
& \\
& \left.+2 N\left(7 m^{2}+6 m-7\right)-8(m-2)\right] \\
& \\
& \left.\quad-(N-2) N^{2}(m+1)(N m+2)(N m+N-2)^{2}\right\},
\end{aligned}
$$

$$
\begin{aligned}
& \text { denom }=4\left\{-D^{8}(m-1)^{4} m\left(N(m-1)+m^{2}-3 m+4\right)\right. \\
& -D^{6}(m-1)^{3}\left[N^{3} m\left(m^{3}-m^{2}-3 m-1\right)+N^{2}\left(-5 m^{4}-5 m^{3}+11 m^{2}+9 m+2\right)\right. \\
& \left.+4 N\left(4 m^{3}-2 m^{2}-4 m-1\right)-2 m\left(m^{2}+6 m-9\right)\right] \\
& +D^{4}(m-1)^{2}\left[\left(N^{5} m(m+1)^{2}\left(3 m^{2}-1\right)-N^{4}\left(9 m^{5}+36 m^{4}+30 m^{3}+m+4\right)\right.\right. \\
& +N^{3}\left(46 m^{4}+96 m^{3}+28 m^{2}+8 m+22\right)-4 N^{2}\left(m^{4}+22 m^{3}+18 m^{2}+11\right) \\
& \left.+8 N\left(m^{3}+9 m^{2}-2 m+4\right)-8 m(m+1)\right] \\
& -D^{2} N(m-1)(N m+N-2)^{2}\left[N^{4} m\left(3 m^{3}+5 m^{2}+3 m+1\right)-N^{3}\left(7 m^{4}+19 m^{3}+15 m^{2}+5 m-2\right)\right. \\
& \left.+4 N^{2}\left(4 m^{3}+6 m^{2}+3 m-2\right)-2 N\left(m^{3}+4 m^{2}+5 m-8\right)+8(m-2)\right] \\
& \left.+(N-2) N^{2}(m+1)(N m+N-2)^{4}\left(N^{2} m^{2}+N m-2\right)\right\} .
\end{aligned}
$$

Above, $N=N_{1}+N_{2}$ is the total population size and $D=\left|N_{1}-N_{2}\right|$ is the difference in size between the islands.

For rare migration $(m \rightarrow 0)$, the critical $b / c$ ratio becomes

$$
\begin{equation*}
\lim _{m \rightarrow 0}\binom{b}{c}^{*}=\frac{N^{2}(N-2)^{3}+N\left(N^{2}-7 N+8\right) D^{2}+(2 N-3) D^{4}}{4 N(N-2)\left(N-D^{2}\right)} \tag{68}
\end{equation*}
$$

If, in addition, the islands are evenly sized $(D=0)$, we have

$$
\lim _{m \rightarrow 0}\left(\frac{b}{c}\right)^{*}=(N-2)^{2} / 4
$$

We can show that the above value is at least a local infimum of $(b / c)^{*}$. This is based on the following observations (made with the aid of Mathematica):
(i) For all $N \geq 4$ and $0<m<1$,

$$
\left.\frac{d(b / c)^{*}}{d D}\right|_{D=0}=0
$$

(ii) For all $N \geq 4$ and $0<m<1$,

$$
\left.\frac{d^{2}(b / c)^{*}}{d D^{2}}\right|_{D=0}>0
$$

(iii) For $D=0$ and all $N \geq 4,(b / c)^{*}$ approaches an infimum as $m \rightarrow 0$.

We conjecture that $(b / c)^{*}=(N-2)^{2} / 4$ is a global infimum as well.

## K.1.2 More than two islands

For three and four islands, we have obtained the exact critical $b / c$ ratio under the limit $m \rightarrow 0$, but the results are too lengthy to record here. For these cases, we have shown with the aid of Mathematica that $\lim _{m \rightarrow 0}(b / c)^{*}$ is minimized when the islands are evenly sized.

For five islands, we have obtained $\lim _{m \rightarrow 0}(b / c)^{*}$ under the assumption that two of the islands have equal size. Under this assumption, we have then shown that $\lim _{m \rightarrow 0}(b / c)^{*}$ is minimized when the remaining three islands also have the same size.

For any number $n$ of evenly-sized islands $\left(N_{i}=N / n\right)$, with arbitrary migration, we have

$$
\begin{equation*}
\binom{b}{c}^{*}=\frac{(N-2)(N-n+m N(n-1))^{2}}{N n\left(N-n+m^{2}(N-1)\right)-2(N-n+m N(n-1))^{2}} . \tag{69}
\end{equation*}
$$

This value of $(b / c)^{*}$ is minimized as $m \rightarrow 0$, at which point it approaches the value

$$
\begin{equation*}
\lim _{m \rightarrow 0}\left(\frac{b}{c}\right)^{*}=\frac{(N-2)(N-n)}{N n-2 N+2 n} . \tag{70}
\end{equation*}
$$

We note that this limiting value is approached in the regime $m \ll 1 / n$. We conjecture that, for each fixed population size $N \geq 4$ and number of islands $n \geq 2$, the right-hand side of Eq. (70) is a global infimum of all positive values of $(b / c)^{*}$ for all migration rates $0<m<1$ and all distributions of island sizes $\left\{N_{i}\right\}_{i=1}^{n}$, as long as there are at least two individuals per island ( $N_{i} \geq 2$ for each $i$ ).

## K. 2 Joined stars

Here we derive the results shown in Fig. 4 bcf of the main text.

Two stars joined by hubs For two $n$-stars joined by an edge between their hubs, solving Eqs. (30) and (31) yield

$$
\begin{equation*}
\tau_{\mathrm{HH}}^{+}=\frac{8 n^{3}+22 n^{2}+17 n+15}{(n+1)^{2}(2 n+5)} \quad \tau_{\mathrm{LL}}^{+}=\frac{12 n+10}{2 n+5} \tag{71}
\end{equation*}
$$

By inspection, we have

$$
\pi_{\mathrm{H}}=\frac{n+1}{4 n+2}, \quad \pi_{\mathrm{L}}=\frac{1}{4 n+2}, \quad p_{\mathrm{HH}}^{(2)}=\frac{n^{2}+n+1}{(n+1)^{2}}, \quad p_{\mathrm{LL}}^{(2)}=\frac{1}{n+1}
$$

Substituting these values into Eq. (41) yields

$$
\binom{b}{c}^{*}=\frac{(n+1)^{2}\left(10 n^{2}+17 n+5\right)}{n\left(4 n^{3}+12 n^{2}+11 n+5\right)} \xrightarrow{n \rightarrow \infty} \frac{5}{2} .
$$

Two stars joined leaf to hub For two $n$-stars with an edge joining the leaf to the hub of the other, a similar procedure yields

$$
\left(\frac{b}{c}\right)^{*}=\frac{(n+1)\left(36 n^{2}+90 n+19\right)}{4 n\left(3 n^{2}+11 n+9\right)} \xrightarrow{n \rightarrow \infty} 3
$$

Two stars joined leaf to leaf For two $n$-stars with an edge joining a leaf of each, we obtain

$$
\left(\frac{b}{c}\right)^{*}=\frac{2(n+1)\left(490 n^{4}+3065 n^{3}+5982 n^{2}+4559 n+1136\right)}{350 n^{5}+2671 n^{4}+6818 n^{3}+7489 n^{2}+3544 n+568} \xrightarrow{n \rightarrow \infty} \frac{14}{5} .
$$

"Dense cluster" of stars For $m$-stars, with each hub joined to each other hub, we obtain

$$
\left(\frac{b}{c}\right)^{*}=\frac{\text { num }}{\text { denom }}
$$

where

$$
\begin{aligned}
& \text { num }=(m+n-1)^{2} \\
& \quad \times\left(2 m^{4}+m^{3}(11 n-8)+m^{2}\left(20 n^{2}-25 n+11\right)\right. \\
& \left.\quad+m\left(12 n^{3}-22 n^{2}+12 n-7\right)-2\left(2 n^{3}+n^{2}+n-1\right)\right), \\
& \begin{array}{r}
\text { denom }=2 m^{5}(n-1)+m^{4}\left(12 n^{2}-19 n+10\right)+m^{3}\left(26 n^{3}-61 n^{2}+61 n-19\right)
\end{array} \\
& +3 m^{2}\left(8 n^{4}-28 n^{3}+40 n^{2}-29 n+6\right)+m\left(8 n^{5}-48 n^{4}+92 n^{3}-103 n^{2}+57 n-9\right) \\
& \quad-8 n^{5}+24 n^{4}-34 n^{3}+32 n^{2}-14 n+2 .
\end{aligned}
$$

Letting $n \rightarrow \infty$ for fixed $m$, we obtain

$$
\lim _{n \rightarrow \infty}\left(\frac{b}{c}\right)^{*}=\frac{3 m-1}{2 m-2}
$$

If we then take $m \rightarrow \infty$, we have $(b / c)^{*} \rightarrow 3 / 2$.

## K. 3 Ceiling fan

Let us now consider the "ceiling fan" graph (Fig. 4e of the main text), in which each of the $n$ leaves of a star is joined by an edge to one other. Solving Eqs. (30) and (31) yields

$$
\tau_{\mathrm{HH}}^{+}=\frac{9 n-3}{n+3}, \quad \tau_{\mathrm{LL}}^{+}=\frac{15 n}{2(n+3)}
$$

where H and L indicate hub and leaf vertices, respectively. By inspection, we have

$$
\pi_{\mathrm{H}}=\frac{1}{3}, \quad \pi_{\mathrm{L}}=\frac{2}{3 n}, \quad p_{\mathrm{HH}}^{(2)}=\frac{1}{2}, \quad p_{\mathrm{LL}}^{(2)}=\frac{n+2}{4 n} .
$$

Substituting these values into Eq. (41) yields

$$
\left(\frac{b}{c}\right)^{*}=\frac{4(6 n-7)}{3 n-16} \xrightarrow{n \rightarrow \infty} 8
$$

## K. 4 Wheel

In a wheel graph (Fig. 4e of the main text), each of the $n$ leaves is joined to two neighboring leaves as well as to the hub. We define $\tau_{L, j}$ to be the coalescence time for two leaves that are $j$ apart, $0 \leq j \leq n$. Clearly, we have $\tau_{L, 0}=\tau_{L, n}=0$. We also define $\tau_{L H}$ to be the coalescence time between a leaf and the hub.

The recurrence relations (30) for coalescence times become

$$
\begin{align*}
\tau_{L, j} & =1+\frac{1}{3}\left(\tau_{L, j-1}+\tau_{L, j+1}+\tau_{L H}\right) \quad \text { for } 1 \leq j \leq n-1  \tag{72}\\
\tau_{L H} & =1+\frac{1}{3} \tau_{L H}+\frac{1}{2 n} \sum_{j=0}^{n-1} \tau_{L, j} . \tag{73}
\end{align*}
$$

Solving Eq. (73) for $\tau_{L H}$ yields

$$
\begin{equation*}
\tau_{L H}=\frac{3}{2}+\frac{3}{4 n} \sum_{j=0}^{n-1} \tau_{L, j} \tag{74}
\end{equation*}
$$

It is convenient to define $\tau_{L, j}^{\prime}=\tau_{L, j}-\tau_{L H}$. Then Eqs. (72) and (74) become

$$
\begin{align*}
\tau_{L, j}^{\prime} & =1+\frac{1}{3}\left(\tau_{L, j-1}^{\prime}+\tau_{L, j+1}^{\prime}\right) \quad \text { for } 1 \leq j \leq n-1  \tag{75}\\
\tau_{L H} & =6+\frac{3}{n} \sum_{j=0}^{n-1} \tau_{L, j}^{\prime} . \tag{76}
\end{align*}
$$

We guess a solution (ansatz) of the form

$$
\begin{equation*}
\tau_{L, j}^{\prime}=a+b\left(\gamma^{j}+\gamma^{n-j}\right) . \tag{77}
\end{equation*}
$$

Substituting this ansatz into Eq. (75), we obtain

$$
\begin{aligned}
a+b\left(\gamma^{j}+\gamma^{n-j}\right) & =1+\frac{2 a}{3}+\frac{b}{3}\left(\gamma^{j-1}+\gamma^{n-j+1}+\gamma^{j+1}+\gamma^{n-j-1}\right) \\
& =1+\frac{2 a}{3}+\frac{b}{3}\left(\gamma+\gamma^{-1}\right)\left(\gamma^{j}+\gamma^{n-j}\right) .
\end{aligned}
$$

For this to hold for all $1 \leq j \leq n-1$ necessitates that

$$
a=1+\frac{2 a}{3} \quad \text { and } \quad \gamma+\gamma^{-1}=3
$$

which gives the solutions

$$
a=3 \quad \text { and } \quad \gamma=\frac{3 \pm \sqrt{5}}{2}
$$

It turns out not to matter which value of $\gamma$ is used; we will use $\gamma=(3-\sqrt{5}) / 2$. To solve for $b$, we substitute into Eq. (76),

$$
\begin{align*}
\tau_{L H} & =6+\frac{3}{n} \sum_{j=0}^{n-1}\left(3+b\left(\gamma^{j}+\gamma^{n-j}\right)\right) \\
& =15+\frac{3 b}{n} \frac{(1+\gamma)\left(1-\gamma^{n}\right)}{1-\gamma} . \tag{78}
\end{align*}
$$

Additionally, since $\tau_{L, 0}=0$, we have

$$
\begin{equation*}
\tau_{L H}=-\tau_{L, 0}^{\prime}=-b\left(1+\gamma^{n}\right)-3 \tag{79}
\end{equation*}
$$

Combining Eqs. (78) and (79) and solving for $b$ yields

$$
b=\frac{-18 n(1-\gamma)}{3(1+\gamma)\left(1-\gamma^{n}\right)+n\left(1+\gamma^{n}\right)(1-\gamma)} .
$$

Substituting this value of $b$ into Eqs. (79) and (77), we obtain the coalescence times

$$
\begin{aligned}
\tau_{L H} & =\frac{18 n(1-\gamma)\left(1+\gamma^{n}\right)}{3(1+\gamma)\left(1-\gamma^{n}\right)+n\left(1+\gamma^{n}\right)(1-\gamma)}-3, \\
\tau_{L, j} & =\tau_{L, j}^{\prime}+\tau_{L H} \\
& =\frac{18 n(1-\gamma)}{3(1+\gamma)\left(1-\gamma^{n}\right)+n\left(1+\gamma^{n}\right)(1-\gamma)}\left(1+\gamma^{n}-\gamma^{j}-\gamma^{n-j}\right) \\
& =\frac{18 n(1-\gamma)}{3(1+\gamma)\left(1-\gamma^{n}\right)+n\left(1+\gamma^{n}\right)(1-\gamma)}\left(1-\gamma^{j}\right)\left(1-\gamma^{n-j}\right)
\end{aligned}
$$

In particular, for neighboring leaves $(j=1)$, we have

$$
\begin{equation*}
\tau_{L, 1}=\frac{18 n(1-\gamma)^{2}\left(1-\gamma^{n-1}\right)}{3(1+\gamma)\left(1-\gamma^{n}\right)+n\left(1+\gamma^{n}\right)(1-\gamma)} \tag{80}
\end{equation*}
$$

Turning now to remeeting times, we compute

$$
\begin{aligned}
\tau_{H H}^{+} & =1+\tau_{L H} \\
& =\frac{18 n(1-\gamma)\left(1+\gamma^{n}\right)}{3(1+\gamma)\left(1-\gamma^{n}\right)+n\left(1+\gamma^{n}\right)(1-\gamma)}-2 \\
\tau_{L L}^{+} & =1+\frac{1}{3} \tau_{L H}+\frac{2}{3} \tau_{L, 1} \\
& =\frac{18 n(1-\gamma)}{3(1+\gamma)\left(1-\gamma^{n}\right)+n\left(1+\gamma^{n}\right)(1-\gamma)}\left(1+\gamma^{n}-\frac{2 \gamma}{3}\left(1+\gamma^{n-2}\right)\right)
\end{aligned}
$$

The other values needed to compute $(b / c)^{*}$ are

$$
\pi_{H}=\frac{1}{4}, \quad \pi_{L}=\frac{3}{4 n}, \quad p_{H H}^{(2)}=\frac{1}{3}, \quad p_{L L}^{(2)}=\frac{2 n+3}{9 n} .
$$

Using the above values, the critical $b / c$ ratio can be obtained from Eq. (41)

$$
\begin{equation*}
\left(\frac{b}{c}\right)^{*}=\frac{\pi_{H} \tau_{H H}^{+}+n \pi_{L} \tau_{L L}^{+}-2}{\pi_{H} \tau_{H H}^{+} p_{H H}^{(2)}+n \pi_{L} \tau_{L L}^{+} p_{L L}^{(2)}-2} . \tag{81}
\end{equation*}
$$

Now turning to the $n \rightarrow \infty$ limit, we calculate:

$$
\begin{aligned}
\lim _{n \rightarrow \infty} \tau_{H H}^{+} & =16 \\
\lim _{n \rightarrow \infty} \tau_{L L}^{+} & =18-12 \gamma=6 \sqrt{5} \\
\lim _{n \rightarrow \infty} p_{L L}^{(2)} & =\frac{2}{9}
\end{aligned}
$$

Substituting into Eq. (81) and simplifying gives

$$
\lim _{n \rightarrow \infty}\left(\frac{b}{c}\right)^{*}=\frac{429+90 \sqrt{5}}{82}
$$

## L Direct and inclusive fitness

The conditions for success derived in Sections $B$ are based on fixation probability and, in the case of nonzero mutation, expected degree-weighted abundance. Other approaches in the literature are based on the fitness and/or inclusive fitness of individuals. In the interest of synthesizing different approaches, we calculate the fitness and -in the case of the donation game (8) - the inclusive fitness effect associated to each vertex.

## L. 1 Fitness

The (direct) fitness of an individual is a measure of its reproductive success. In homogeneous populations, the fitness of an individual is defined as its survival probability plus its expected number of offspring. For heterogeneous populations, individuals have different reproductive values-that is, they make different expected contributions to the future gene pool of the population, even under neutral drift [52, 53, 65, [58]. For our model, we identify the reproductive value of vertex $i$ as its relative weighted degree $\pi_{i}$ [53], which is also equal to the fixation probability of a neutral mutation arising at this vertex [53, 19 .

We formally define the fitness of an individual as its survival probability multiplied by its own reproductive value, plus the expected total reproductive value of all offspring it produces, over a short time interval $[t, t+\epsilon)$. The
fitness $v_{i}(\mathbf{s})$ of vertex $i$ in state $\mathbf{s}$ is calculated as follows:

$$
\begin{align*}
v_{i}(\mathbf{s}) & =\pi_{i}\left(1-\epsilon \sum_{j \in G} \operatorname{Rate}[j \rightarrow i](\mathbf{s})\right)+\epsilon \sum_{j \in G} \pi_{j} \operatorname{Rate}[i \rightarrow j](\mathbf{s})+o(\epsilon) \\
& =\pi_{i}+\epsilon\left(\sum_{j \in G} \frac{\pi_{j} w_{i j} F_{i}(\mathbf{s})}{\sum_{k \in G} w_{k j} F_{k}(\mathbf{s})}-\pi_{i}\right)+o(\epsilon) \\
& =\pi_{i}+\epsilon \delta \pi_{i}\left(f_{i}(\mathbf{s})-f_{j}^{(2)}(\mathbf{s})\right)+R(\epsilon, \delta) \tag{82}
\end{align*}
$$

where the remainder term $R(\epsilon, \delta)$ satisfies

$$
\lim _{\epsilon \rightarrow 0^{+}} \lim _{\delta \rightarrow 0} \frac{R(\epsilon, \delta)}{\epsilon \delta}=0
$$

The constant term in Eq. (82) is the fitness of vertex $i$ under neutral drift, which is equal to its reproductive value $\pi_{i}$. The second term represents the effects of weak selection. We define the direct fitness effect of selection on individual $i$, denoted $v_{i}^{\prime}(\mathbf{s})$ to be the coefficient of $\epsilon \delta$ :

$$
\begin{equation*}
v_{i}^{\prime}(\mathbf{s})=\pi_{i}\left(f_{i}(\mathbf{s})-f_{i}^{(2)}(\mathbf{s})\right) \tag{83}
\end{equation*}
$$

For an arbitrary matrix game (6), substituting the payoffs from Eq. (7), we obtain

$$
\begin{align*}
v_{i}^{\prime}(\mathbf{s})=\pi_{i}(a & {\left[s_{i} s_{i}^{(1)}-\sum_{j \in G} p_{i j}^{(2)} s_{j} s_{j}^{(1)}\right] } \\
& +b\left[s_{i}\left(1-s_{i}^{(1)}\right)-\sum_{j \in G} p_{i j}^{(2)} s_{j}\left(1-s_{j}^{(1)}\right)\right] \\
& +c\left[\left(1-s_{i}\right) s_{i}^{(1)}-\sum_{j \in G} p_{i j}^{(2)}\left(1-s_{j}\right) s_{j}^{(1)}\right]  \tag{84}\\
& \left.+d\left[\left(1-s_{i}\right)\left(1-s_{i}^{(1)}\right)-\sum_{j \in G} p_{i j}^{(2)}\left(1-s_{j}\right)\left(1-s_{j}^{(1)}\right)\right]\right) .
\end{align*}
$$

For the donation game (8), Eq. (84) simplifies to

$$
\begin{equation*}
v_{i}^{\prime}(\mathbf{s})=\pi_{i}\left(-c s_{i}+b s_{i}^{(1)}+c s_{i}^{(2)}-b s_{i}^{(3)}\right) \tag{85}
\end{equation*}
$$

Eqs. (83)-(85) apply to a particular state s. We can also calculate the direct fitness effect of a particular strategy at a particular vertex for the overall evolutionary process. For this we introduce mutation with probability $0<u<1$, as discussed in Section I. The expected direct fitness effect of strategy A at vertex $i$, under the neutral MSS distribution, can be written as

$$
\begin{equation*}
\mathbb{E}_{\mathrm{MSS}}^{\circ}\left[v_{i}^{\prime} \mid s_{i}=1\right]=\pi_{i} \mathbb{E}_{\mathrm{MSS}}^{\circ}\left[f_{i}-f_{i}^{(2)} \mid s_{i}=1\right] \tag{86}
\end{equation*}
$$

To explicitly compute the right-hand side of (86) for the general game (6) requires triplet IBD probabilities [66, 67] and is beyond the scope of this work. However, for the donation game (8), only pairwise IBD probabilities are required. A generalization of Lemma 3 of [6] yields

$$
\mathbb{E}_{\mathrm{MSS}}^{\circ}\left[s_{j} \mid s_{i}=1\right]=\frac{1+q_{i j}}{2}
$$

Applying this identity to Eq. (85) yields the overall direct fitness effect of cooperation at vertex $i$ :

$$
\begin{aligned}
\mathbb{E}_{\mathrm{MSS}}^{\circ}\left[v_{i}^{\prime} \mid s_{i}=1\right] & =\frac{\pi_{i}}{2}\left(-c\left(q_{i}^{(0)}-q_{i}^{(2)}\right)+b\left(q_{i}^{(1)}-q_{i}^{(3)}\right)\right) \\
& =\frac{u \pi_{i}}{2}\left[-c \tau_{i}^{(2)}+b\left(\tau_{i}^{(3)}-\tau_{i}^{(1)}\right)\right]+\mathcal{O}\left(u^{2}\right) \quad(u \rightarrow 0)
\end{aligned}
$$

Above, we have defined $q_{i}^{(n)}=\sum_{j \in G} p_{i j}^{(n)} q_{i j}$ and $\tau_{i}^{(n)}=\sum_{j \in G} p_{i j}^{(n)} \tau_{i j}$. The second equality above uses Eq. (60).

## L. 2 Inclusive fitness

Inclusive fitness theory [68, 56, 4] analyzes the evolution of social behavior using a quantity called the inclusive fitness effect. This quantity is defined as the effect that this individual has on its own fitness, plus a weighted sum of the effects it has on the fitnesses of all others, where the weights quantify genetic relatedness.

In order to formulate such a quantity, there must be a well-defined contribution that each individual makes to its own fitness and to the fitness of each other individual. However, Eq. (84) for the direct fitness effect is quadratic in $s_{1}, \ldots, s_{N}$, and does not separate into distinct contributions due to particular individuals. Therefore, for the general game (6), there is no inclusive fitness effect of a single given individual. (But see Refs. [66, 67, 69] for alternative notions of inclusive fitness at the level of pairs or genetic lineages.)

For the donation game (8) the direct fitness effect (85) is linear in $s_{1}, \ldots, s_{N}$ and does separate into distinct contributions due to particular individuals. Specifically, we can write

$$
\begin{equation*}
v_{i}^{\prime}(\mathbf{s})=\sum_{j \in G} e_{j i} s_{j} \tag{87}
\end{equation*}
$$

where we have defined the fitness effect of cooperation at $j$ on individual $i$ as

$$
\begin{equation*}
e_{j i}=\pi_{i}\left(-c p_{i j}^{(0)}+b p_{i j}^{(1)}+c p_{i j}^{(2)}-b p_{i j}^{(3)}\right) . \tag{88}
\end{equation*}
$$

Because the fitness effects $e_{i j}$ are well-defined for game (8), the inclusive fitness effect of cooperation at vertex $i$ exists and can be written as

$$
\begin{equation*}
v_{i}^{\mathrm{IF}}(\mathbf{s})=\sum_{j \in G} e_{i j} s_{j} . \tag{89}
\end{equation*}
$$

In the above expression, the relatedness of a hypothetical cooperator at vertex $i$ to the occupant of vertex $j$ is defined to be $s_{j}$. This notion of relatedness applies to a particular state $\mathbf{s}$, and might be termed "identity in state". It is common to normalize relatedness coefficients so that they lie in the range -1 to 1 ; for example, one might use $\left(s_{j}-\bar{s}\right) /(1-\bar{s})$ instead of $s_{j}$ to quantify the relatedness of a cooperator at $i$ to the occupant of $j$, where $\bar{s}=\frac{1}{N} \sum_{k \in G} s_{k}$ is the average population type. However, such normalizations are not needed for models with constant population size [4].

We observe that, as a consequence of the reversibility property $\pi_{i} p_{i j}^{(n)}=$ $\pi_{j} p_{i j}^{(n)}$, fitness effects are symmetric: $e_{i j}=e_{j i}$. In other words, the effect that cooperation at $i$ has on vertex $j$ is equal to the effect that cooperation at $j$ has on vertex $i$. It follows that the direct and inclusive fitness effects of cooperation at $i$ are equal in every state: $v_{i}^{\mathrm{IF}}(\mathbf{s})=v_{i}^{\prime}(\mathbf{s})$. This is an interesting but idiosyncratic property of the model we consider. We would not, for example, find the result for different interaction and replacement graphs (Section J.2) since there is no analogous reversibility property for ( $n, m$ )-random walks.

Since - for the donation game (8) - the direct and inclusive fitness effects of cooperation at $i$ are equal in every state, they are also equal for the overall
evolutionary process:

$$
\begin{aligned}
\mathbb{E}_{\mathrm{MSS}}^{\circ}\left[v_{i}^{\mathrm{IF}} \mid s_{i}=1\right] & =\mathbb{E}_{\mathrm{MSS}}^{\circ}\left[v_{i}^{\prime} \mid s_{i}=1\right] \\
& =\frac{\pi_{i}}{2}\left(-c\left(q_{i}^{(0)}-q_{i}^{(2)}\right)+b\left(q_{i}^{(1)}-q_{i}^{(3)}\right)\right) \\
& =\frac{u \pi_{i}}{2}\left[-c \tau_{i}^{(2)}+b\left(\tau_{i}^{(3)}-\tau_{i}^{(1)}\right)\right]+\mathcal{O}\left(u^{2}\right) \quad(u \rightarrow 0)
\end{aligned}
$$

## M Computational issues

As we pointed out below Eq. (30), computing the coalescence times involves solving a system of $\binom{N}{2}$ linear equations. Simple Gaussian elimination takes $\mathcal{O}\left(N^{6}\right)$ steps for such a system. However this can be improved by a (standard) block-wise inversion approach combined with a state-of-the-art matrix multiplication algorithm. For example, based on variants of the CoppersmithWinograd algorithm, coalescence times can be computed in $\mathcal{O}\left(N^{4.75}\right)$ time.

Further improvement can be achieved by allowing approximate solutions and observing that (30) is a symmetric diagonally dominant (SDD) system. Such systems can be solved in nearly linear time in the number of non-zero entries of the coefficient matrix. More precisely, for a system described by an $n \times n$ matrix with $m$ non-zero entries, finding a vector $\epsilon$ far in norm from the exact solution can be done in $\mathcal{O}\left(m \log ^{2} n \log 1 / \epsilon\right)$ time [70]. It follows that for a graph of size $N$ with average degree $\bar{k}$, coalescence times can be determined to within $\epsilon$ in $\mathcal{O}\left((N \log N)^{2} \bar{k} \log 1 / \epsilon\right)$ time. Furthermore, the algorithm can be efficiently parallelized [71].

In our experiments we have used a MatLab implementation by Koutis (http://tiny.cc/cmgSolver). Representative run times on a 2011 MacBook Air, MatLab 2015a were as follows: for $N=1000$, average degree 4, the running time was 12 seconds. For $N=2000$, average degree 4 , the running time was 120 seconds. For $N=2000$, average degree 8, the running time was 280 seconds.

For Figure 5 we computed $(b / c)^{*}$ for 1.3 million unweighted graphs, generated from 10 different random graph models. Parameter values were sampled from a uniform distribution on the specified ranges (see below). Initial graph sizes were uniformly sampled in the range $100 \leq N \leq 150$; if the random graph model produced a disconnected graph, the largest connected component was used. The critical $(b / c)^{*}$ ratio was computed by solving Eq. (2) for
coalescence times and substituting into Eq. (41). (No Monte Carlo simulations were used for these investigations.)

Random graph models and parameter ranges are as follows: 100K ErdosRenyi [39] with edge probability $0<p<1$; 100K small world [32] with initial connection distance $1 \leq d \leq 5$ and edge creation probability $0<p<0.05$; 100K Barabasi-Albert [31] with linking number $1 \leq m \leq 10$; 100K random recursive [35] (like Barabasi-Albert except that edges are added uniformly instead of preferentially) with linking number $2 \leq m \leq 8$, 200K HolmeKim [34] with linking number $2 \leq m \leq 4$ and triad formation parameter $0<P<0.15 ; 200 \mathrm{~K}$ Klemm-Eguiluz [33] with linking number $3 \leq m \leq 5$ and deactivation parameter $0<\mu<0.15$; 200K shifted-linear preferential attachment [36] with linking number $1 \leq m \leq 7$ and shift $0<\theta<40 ; 100 \mathrm{~K}$ forest fire [37] with parameters $0<p_{f}<p_{b}<0.15 ; 100 \mathrm{~K}$ Island BA; and 100K Island ER. Island BA is a meta-network of islands [38], in which each island is a shifted-linear preferential attachment with the same parameters as above. The number of islands varies from 2 to 5 . Considering the islands as meta-nodes, the meta-network among the islands is an ER graph with edge probability $0<p_{\text {inter }}<1$. Island ER is the as Island BA except that each island is an ER graph with edge probability $0<p_{\text {intra }}<1$.

We also computed the critical $b / c$ ratio for some large real-world networks using Northeastern's computational cluster. This was computed on a node with Intel Xeon CPU E5-2680 2.8GHz and 256GB RAM.

- The Framingham Study graph $(N=5253$, average degree 6.5$)$. This network has a critical $b / c$ ratio of 7.96. The running time was around 3.5 hours.
- The ego-facebook network from the Stanford SNAP database ( $N=$ 4039, average degree 43.7) took 25 minutes and has critical $b / c$ ratio 48.5.
- The (largest connected component of the) ca-GrQc graph from the Stanford SNAP database ( $N=4159$, average degree 6.5 took 23 hours and has critical $b / c$ ratio 6.6. The reason for this is most likely that this graph is very badly conditioned.


Figure 12: Results from additional Monte Carlo simulations. Simulation are shown for averaged payoffs (solid dots) and accumulated payoffs (open dots). Vertical lines indicate theoretical $(b / c)^{*}$ for averaged and accumulated payoffs, as well as the mean degree $\bar{k}$. The horizontal line pertains to neutral drift, for which $N \rho=1$. (a) Barabasi-Albert network [31] with $m=4, N=100$. (b) Klemm-Eguiluz network [33] with $m=3, \mu=0.4, N=100$. (c) Small world network [32] with initial connection distance $d=3$, link creation probability $p=0.1$ and $N=80$. (d) Forest fire network [37] with parameters $p_{b}=0.32, p_{f}=0.28, N=80$.

## N Monte Carlo Simulations

To verify that our results (which are exact in the limit of weak selection) are accurate for nonweak selection, we performed Monte Carlo simulations. Results are presented in Figures 2 and 12.

The simulation setup is as follows. For each graph, we run $5 \times 10^{5}$ Monte Carlo trials. For each trial, all nodes are defectors upon the inception, except one randomly-selected node which is a cooperator. The fixation probability is approximated as the fraction of Monte Carlo trials which eventuate in unanimous cooperation before timestep $T$, which is set to 400000 (which theoretically should be infinite). The cost of cooperation is $c=1$. For Fig. 2, selection strength is $\delta=0.025$. For Fig. 12, selection strength is $\delta=10^{-2}$ for averaged payoffs and $\delta=10^{-3}$ for accumulated payoffs. The smaller $\delta$ value for accumulated payoffs is needed to compensate for the summing of a potentially large number of individual payoffs.

## References

[1] Erez Lieberman, Christoph Hauert, and Martin A Nowak. Evolutionary dynamics on graphs. Nature, 433(7023):312-316, 2005.
[2] M. A. Nowak, C .E. Tarnita, and T. Antal. Evolutionary dynamics in structured populations. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1537):19-30, 2010.
[3] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. Nature, 441:502-505, 2006.
[4] P. D. Taylor, T. Day, and G. Wild. Evolution of cooperation in a finite homogeneous graph. Nature, 447(7143):469-472, 2007.
[5] Yu-Ting Chen. Sharp benefit-to-cost rules for the evolution of cooperation on regular graphs. The Annals of Applied Probability, 23(2):637664, 2013.
[6] Benjamin Allen and Martin A Nowak. Games on graphs. EMS Surveys in Mathematical Sciences, 1(1):113-151, 2014.
[7] F Débarre, C Hauert, and M Doebeli. Social evolution in structured populations. Nature Communications, 5:4409, 2014.
[8] Rasmus Ibsen-Jensen, Krishnendu Chatterjee, and Martin A Nowak. Computational complexity of ecological and evolutionary spatial dynamics. Proceedings of the National Academy of Sciences, 112(51):1563615641, 2015.
[9] John Frank Charles Kingman. The coalescent. Stochastic Processes and their Applications, 13(3):235-248, 1982.
[10] John Wakeley. Coalescent Theory: An Introduction. Roberts \& Company Publishers, Greenwood Village, CO, 2009.
[11] J Theodore Cox. Coalescing random walks and voter model consensus times on the torus in $\mathbf{Z}^{d}$. Annals of Probability, 17(4):1333-1366, 1989.
[12] R. Durrett and S. Levin. The importance of being discrete (and spatial). Theoretical Population Biology, 46(3):363-394, 1994.
[13] M. P. Hassell, H. N. Comins, and R. M. May. Species coexistence and self-organizing spatial dynamics. Nature, 370:290-292, 1994.
[14] M. A. Nowak and R. M. May. Evolutionary games and spatial chaos. Nature, 359(6398):826-829, 1992.
[15] Benjamin Allen, Jeff Gore, and Martin A Nowak. Spatial dilemmas of diffusible public goods. eLife, 2, 2013.
[16] Martin A. Nowak, Franziska Michor, and Yoh Iwasa. The linear process of somatic evolution. Proceedings of the National Academy of Sciences, 100(25):14966-14969, 2003.
[17] Francisco C Santos, Marta D Santos, and Jorge M Pacheco. Social diversity promotes the emergence of cooperation in public goods games. Nature, 454(7201):213-216, 2008.
[18] David G Rand, Martin A Nowak, James H Fowler, and Nicholas A Christakis. Static network structure can stabilize human cooperation. Proceedings of the National Academy of Sciences, 111(48):17093-17098, 2014.
[19] Benjamin Allen, Christine Sample, Yulia Dementieva, Ruben C. Medeiros, Christopher Paoletti, and Martin A. Nowak. The molecular clock of neutral evolution can be accelerated or slowed by asymmetric spatial structure. PLoS Computational Biology, 11(2):e1004108, 02 2015.
[20] J. Maynard Smith. Evolution and the Theory of Games. Cambridge University Press, Cambridge, 1982.
[21] J. Hofbauer and K. Sigmund. Evolutionary Games and Replicator Dynamics. Cambridge University Press, Cambridge, UK, 1998.
[22] M Broom and J Rychtár. Game-Theoretical Models in Biology. Chapman \& Hall/CRC, Boca Raton, FL, USA, 2013.
[23] Thomas M Liggett. Interacting Particle Systems. Springer Science \& Business Media, 2006.
[24] F. C. Santos and J. M. Pacheco. Scale-free networks provide a unifying framework for the emergence of cooperation. Physical Review Letters, 95(9):98104, 2005.
[25] Wes Maciejewski, Feng Fu, and Christoph Hauert. Evolutionary game dynamics in populations with heterogenous structures. PLoS computational biology, 10(4):e1003567, 2014.
[26] Steven H Strogatz. Exploring complex networks. Nature, 410(6825):268276, 2001.
[27] Tomohiko Konno. A condition for cooperation in a game on complex networks. Journal of Theoretical Biology, 269(1):224-233, 2011.
[28] Corina E. Tarnita, Hisashi Ohtsuki, Tibor Antal, Feng Fu, and Martin A. Nowak. Strategy selection in structured populations. Journal of Theoretical Biology, 259(3):570-581, 2009.
[29] C Hadjichrysanthou, M Broom, and J Rychtář. Evolutionary games on star graphs under various updating rules. Dynamic Games and Applications, 1(3):386-407, 2011.
[30] J. Gore, H. Youk, and A. Van Oudenaarden. Snowdrift game dynamics and facultative cheating in yeast. Nature, 459(7244):253-256, 2009.
[31] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. science, 286(5439):509-512, 1999.
[32] Mark EJ Newman and Duncan J Watts. Renormalization group analysis of the small-world network model. Physics Letters A, 263(4):341-346, 1999.
[33] Konstantin Klemm and Victor M Eguiluz. Highly clustered scale-free networks. Physical Review E, 65(3):036123, 2002.
[34] Petter Holme and Beom Jun Kim. Growing scale-free networks with tunable clustering. Physical Review E, 65(2):026107, 2002.
[35] Sergey N Dorogovtsev, Alexander V Goltsev, and José FF Mendes. Critical phenomena in complex networks. Reviews of Modern Physics, 80(4):1275, 2008.
[36] Paul L Krapivsky and Sidney Redner. Organization of growing random networks. Physical Review E, 63(6):066123, 2001.
[37] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, pages 177-187. ACM, 2005.
[38] Sergi Lozano, Alex Arenas, and Angel Sánchez. Mesoscopic structure conditions the emergence of cooperation on social networks. PLoS ONE, 3(4):e1892, 2008.
[39] Paul Erdös and Alfréd Rényi. On random graphs, i. Publicationes Mathematicae (Debrecen), 6:290-297, 1959.
[40] David Lusseau, Karsten Schneider, Oliver J Boisseau, Patti Haase, Elisabeth Slooten, and Steve M Dawson. The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. Behavioral Ecology and Sociobiology, 54(4):396-405, 2003.
[41] DS Sade. Sociometrics of macaca mulatta i. linkages and cliques in grooming matrices. Folia primatologica, 18(3-4):196-223, 1972.
[42] Siva R Sundaresan, Ilya R Fischhoff, Jonathan Dushoff, and Daniel I Rubenstein. Network metrics reveal differences in social organization between two fission-fusion species, grevy's zebra and onager. Oecologia, 151(1):140-149, 2007.
[43] Alison L. Hill, David G. Rand, Martin A. Nowak, and Nicholas A. Christakis. Infectious disease modeling of social contagion in networks. PLoS Computational Biology, 6(11):e1000968, November 2010.
[44] Nicholas A Christakis and James H Fowler. The spread of obesity in a large social network over 32 years. New England Journal of Medicine, 357(4):370-379, 2007.
[45] Julian J McAuley and Jure Leskovec. Learning to discover social circles in ego networks. Advances in Neural Information Processing Systems, 25:548-56, 2012.
[46] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. ACM Transactions on Knowledge Discovery from Data (TKDD), 1(1):2, 2007.
[47] Christoph Hauert, Franziska Michor, Martin A Nowak, and Michael Doebeli. Synergy and discounting of cooperation in social dilemmas. Journal of Theoretical Biology, 239(2):195-202, 2006.
[48] Martin A Nowak. Evolving cooperation. Journal of Theoretical Biology, 299:1-8, 2012.
[49] Hisashi Ohtsuki, Martin A. Nowak, and Jorge M. Pacheco. Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs. Physical Review Letters, 98(10):108106, Mar 2007.
[50] Benjamin Allen, Arne Traulsen, Corina E. Tarnita, and Martin A. Nowak. How mutation affects evolutionary games on graphs. Journal of Theoretical Biology, 299:97-105, 2012.
[51] Benjamin Allen and Corina E Tarnita. Measures of success in a class of evolutionary models with fixed population size and structure. Journal of Mathematical Biology, 68(1-2):109-143, 2014.
[52] Peter D Taylor. Allele-frequency change in a class-structured population. American Naturalist, pages 95-106, 1990.
[53] Wes Maciejewski. Reproductive value in graph-structured populations. Journal of theoretical biology, 340:285-293, 2014.
[54] P. D. Taylor and S. A. Frank. How to make a kin selection model. Journal of Theoretical Biology, 180(1):27-37, 1996.
[55] Henri Leturque and François Rousset. Dispersal, kin competition, and the ideal free distribution in a spatially heterogeneous population. Theoretical Population Biology, 62(2):169-180, 2002.
[56] F. Rousset. Genetic Structure and Selection in Subdivided Populations. Princeton University Press, Princeton, 2004.
[57] S. Lessard and V. Ladret. The probability of fixation of a single mutant in an exchangeable selection model. Journal of Mathematical Biology, 54(5):721-744, 2007.
[58] Corina E Tarnita and Peter D Taylor. Measures of relative fitness of social behaviors in finite structured population models. The American Naturalist, 184(4):477-488, 2014.
[59] Jeremy Van Cleve. Social evolution and genetic interactions in the short and long term. Theoretical Population Biology, 103:2-26, 2015.
[60] David Aldous and Jim Fill. Reversible markov chains and random walks on graphs. Unfinished monograph, 2002.
[61] P.D. Taylor, T. Day, and G. Wild. From inclusive fitness to fixation probability in homogeneous structured populations. Journal of Theoretical Biology, 249(1):101-110, 2007.
[62] Gustave Malécot. Les Mathématiques de l'Hérédité. Masson et Cie., Paris, 1948.
[63] Montgomery Slatkin. Inbreeding coefficients and coalescence times. Genetical Research, 58(02):167-175, 1991.
[64] Hisashi Ohtsuki, Jorge M Pacheco, and Martin A Nowak. Evolutionary graph theory: breaking the symmetry between interaction and replacement. Journal of Theoretical Biology, 246(4):681-694, 2007.
[65] Peter D Taylor and Wes Maciejewski. Hamilton's inclusive fitness in finite-structured populations. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1642):20130360, 2014.
[66] Peter Taylor. Inclusive and personal fitness in synergistic evolutionary games on graphs. Journal of Theoretical Biology, 325:76-82, 2013.
[67] Peter Taylor. Hamilton s rule in finite populations with synergistic interactions. Journal of Theoretical Biology, 397:151-157, 2016.
[68] W. D. Hamilton. The genetical evolution of social behaviour. I. Journal of Theoretical Biology, 7(1):1-16, 1964.
[69] Erol Akçay and Jeremy Van Cleve. There is no fitness but fitness, and the lineage is its bearer. Phil. Trans. R. Soc. B, 371(1687):20150085, 2016.
[70] Ioannis Koutis, Gary L. Miller, and Richard Peng. Approaching optimality for solving sdd linear systems. SIAM Journal on Computing, 43(1):337-354, 2014.
[71] Rasmus Kyng, Yin Tat Lee, Richard Peng, Sushant Sachdeva, and Daniel A. Spielman. Sparsified cholesky and multigrid solvers for connection laplacians. In Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, pages 842-850, New York, NY, USA, 2016. ACM.

