Conjoining uncooperative societies facilitates evolution of cooperation

Babak Fotouhi\(^1,2\), Naghmeh Momeni\(^1,3\), Benjamin Allen\(^4,5\) and Martin A. Nowak\(^6,7\)

Social structure affects the emergence and maintenance of cooperation. Here, we study the evolutionary dynamics of cooperation in fragmented societies, and show that conjoining segregated cooperation-inhibiting groups, if done properly, rescues the fate of collective cooperation. We highlight the essential role of intergroup ties, which sew the patches of the social network together and facilitate cooperation. We point out several examples of this phenomenon in actual settings. We explore random and non-random graphs, as well as empirical networks. In many cases, we find a marked reduction of the critical benefit-to-cost ratio needed for sustaining cooperation. Our finding gives hope that the increasing worldwide connectivity, if managed properly, can promote global cooperation.

A core problem in evolutionary game theory is that of cooperation. Cooperation involves individuals paying a cost to benefit others and is a ubiquitous feature of social life\(^1\). The structure of social networks affects pathways of information, exchange and other interpersonal mechanisms that undergird cooperation\(^2\). Thus, a natural question in the mathematical study of evolutionary dynamics of cooperation is how network structure influences collective cooperative outcomes\(^4\).

Here, we look at the evolution of cooperation from a new perspective. We ask the question of how the interconnection between segregated groups can promote cooperation. Similar to individuals forming groups towards collective individually implausible accomplishments, sometimes groups come together to form larger composite structures. Examples abound throughout history, from trade and intermarriage relations between tribes and communities in antiquity, to the waves of globalization that increasingly connect local entities for economic, cultural and technological exchange. Another example is project management at different levels in corporations and organizations, which involves the cooperative division of labour between sparsely interconnected, distinctly specialized units.

We use the framework of evolutionary graph theory\(^4,5,8\) to study settings where groups that are individually undesirable for cooperation can be conjoined to build larger cooperation-promoting structures. We first study the conjoining of cohesive communities (clique-like structurally homogeneous groups) under different connection schemes. We then focus on extremely heterogeneous structures. We study stars and their various interconnection schemes, as well as rich clubs, and introduce ensuing topologies that are super-promoters of cooperation. Then, we focus on bipartite graphs. In addition to these ideal graph families, we consider several random graph models. Finally, we consider empirical social networks and investigate the role of community structure on the evolution of cooperation. The findings are consistent across topologies: sparse interconnections of cooperation-inhibiting graphs lead to composite structures that are better for the evolution of cooperation.

Under the framework of mathematical graph theory, social structure is described by a graph, in which nodes represent individuals and links represent interactions and/or relations. In the simplest setting, individuals are conventionally envisaged with two possible strategies pertaining to a $2 \times 2$ context-specific payoff matrix that characterizes their interaction. The outcomes of these interactions (‘games’) determine the ‘fitness’ values of the individuals: those who accrue more benefits are endowed with higher fitness, which governs their influence over the peers’ choices of strategy. The most stringent form of cooperation is found in the prisoner’s dilemma game, in which individuals are either cooperators (paying a cost $c$ and bestowing benefit $b > c$ on the interaction partner) or defectors (who seek to benefit without paying a cost). The analysis throughout this paper uses the ‘donation game’ version of the prisoner’s dilemma (as shall be discussed, generalization to arbitrary symmetric two-player games is straightforward).

In this game, mutual cooperation has payoff $b-c$, unilateral cooperation has payoff $-c$ for the cooperator and $b$ for the defector, and mutual defection has payoff $0$. The ratio $b/c$ characterizes the trade-off that players face. Throughout this paper, without loss of generality, we set $c=1$. This is simply equivalent to a change of scale in payoffs and helps the brevity of notation. The strategies of the agents change according to death–birth (dB) updating: a random individual is chosen to update; it adopts one of the neighbours’ strategies proportional to payoff. The small ‘$d$’ indicates that death is random, whereas the large ‘$B$’ indicates that birth is under selection. The probability that the chosen node copies the strategy of neighbour $y$ is proportional to $1 + \delta x$, where $\delta$ denotes the selection strength and $x$ is the average payoff that node $y$ gains playing with its own neighbours. We consider the limit of weak selection. To see whether natural selection favours or hinders collective cooperation, we must calculate the probability that a single cooperator emerging at a random place in the network takes over the population. Natural selection favours cooperation if this fixation probability exceeds the fixation probability of a defector. Otherwise, natural selection inhibits cooperation.

Before we proceed, we point out a central feature of network models of cooperation, such as ours. In these models, social influence spreads beyond immediate neighbours. In conventional models of social contagion, such as simple contagion models, which often describe information diffusion, and complex contagion models,
which often describe the spread of behaviours\textsuperscript{1,2}, the ego’s activation probability depends on the states of the alters. An activated alter exerts the same influence on the ego regardless of the states of the neighbours of that alter. For example, in the simple contagion model of information diffusion, the ego needs to have heard the news from only one alter to have become informed and is agnostic to how many neighbours of that alter have heard the news. Similarly in threshold models of complex contagion, the ego is activated once a certain number or fraction of alters are activated, regardless of the ego’s second neighbours. By contrast, owing to the strategic nature of cooperative dynamics, in our model, the radius of influence is two\textsuperscript{41}. Ego is influenced directly by the strategies of the alters (from whom ego copies its strategy) and also indirectly by those of the neighbours of each alter (who contribute to the payoff of that alter). Our model, with a setting similar to the previous theoretical\textsuperscript{24–26} and experimental\textsuperscript{27,28} studies of human cooperation, thereby adds a strategic element to pure imitation dynamics. Our model shares one similarity with simple contagion processes: having one alter who has adopted each of the strategies makes the ego’s adoption probability for that strategy non-zero.

A recently discovered formulation gives the exact condition under which natural selection favours cooperation on a given network\textsuperscript{4}. The solution utilizes the mathematical equivalence of the problem to that of coalescing random walks on the graph, and the solution is in terms of the remeeting times of random walkers initiated at each node. In the Methods section, we provide a brief overview of the framework. For a given network, the framework produces a quantity (which we denote by \( b^* \)) that determines the fate of cooperation. For any network, we have \(| b^*| > 1 \). If \( b^* \) is positive, then \( b^* \) is the critical benefit-to-cost ratio. That is, natural selection promotes cooperation on the given network if the benefit-to-cost ratio is greater than \( b^* \). The closer \( b^* \) is to unity, the better the network is for promoting cooperation. Conversely, if \( b^* \) is negative, natural selection inhibits cooperation for any benefit-to-cost ratio. In these cases, the network promotes ‘spite’ instead of cooperation. That is, individuals are willing to pay a cost to reduce the payoff of others. The closer the value that \( b^* \) is to \(-1\), the more strongly the network promotes spite. The convention of the literature has hitherto been using \( b^* \) to characterize the conduciveness of networks for cooperation\textsuperscript{5,13}. In Supplementary Method 1.2., we remark that \( 1/b^* \) can also be used; we discuss the advantages of each measure and we find that some of the numerical results are visually better presentable using \( 1/b^* \) instead of \( b^* \). For consistency with the previous literature, we use \( b^* \) to present the results in the main text.

We consider distinct settings in which structures that are known to inhibit cooperation can be connected under various schemes to create larger structures that are conducive to cooperation. In the main text, for brevity, we only provide the simplified version of the results in the large-\( n \) limit (that is, the leading term), and present the full expressions in the corresponding Supplementary Methods. We use the terminology of asymptotic analysis throughout. We say that \( b^* \) grows as \( an^p \), denoted \( b^* \sim an^p \), if \( \lim_{n \to \infty} b^*/(an^p) = 1 \).

Equivalent, we call \( an^p \) the leading term of \( b^* \). Here, \( a \) is the prefactor of the leading term, and the exponent \( p \) is usually of main interest.

Suppose there is a complete graph (clique) of \( n \) nodes. For a clique, selection does not favour cooperation, regardless of \( b \). Namely, the value of \( b^* \) that the method gives is negative. This means that cliques promote spite.

In real-world networks, communities can join together to form larger structures that are better for cooperation. Suppose there are two cliques (which, for simplicity, we consider to be of the same size, and the analytical steps for the general case are the same) and we connect a node from the first clique to a node in the second clique (Fig. 1a). We call these two nodes ‘gate nodes’ and the rest of the nodes in the two communities ‘commoners’. In organizational settings, for example, these gate nodes are called ‘boundary spanners’. They are essential for intergroup flow of information and ideas, intergroup coordination and collaboration, and organizational effectiveness and novelty\textsuperscript{14}. For two communities of size \( n \) with the described interconnection, \( b^* \) is positive and finite, but it grows as \( 1 \times n^2 \). Thus, in principle, it is possible that natural selection favours cooperation, but the necessary \( b^* \) grows quickly with network size. This might be infeasible for actual settings. Connecting the gate nodes via an intermediary ‘broker’ node (Fig. 1b) reduces the leading term to \((2/5) \times n^2\), which is slightly better, but it still grows quickly with \( n \). Marked reduction of \( b^* \) ensues if instead of one broker, there are two brokers on the path between the gate nodes (Fig. 1c). Each group is connected to a third-party trustee node, or representative, and exchange is done via these two nodes. With two broker nodes in the middle, the leading term of \( b^* \) drops to \( 4m \); thus, \( b^* \) grows considerably slower with network size. This interconnection scheme offers a substantial improvement, and the two communities, which individually promote spite, can now be conjoined to form a new composite network that supports cooperation with more plausible values of \( b^* \).

Longer chains of intermediary nodes between the two cliques is mathematically possible, but relatively less common in actual settings. The possible exceptions are chain-of-command structures, which resemble this topology: a group of decision makers sit at one end (the first clique) and through a chain of intermediary units, the
agenda reaches the bottom-most unit (the second clique), which is in charge of implementation. For chains with more than two intermediary nodes, the analytical results become too lengthy to be presented. But fortunately, our coalescing random walks framework enables numerical extraction of the leading term. If the chain of intermediaries has length \( L \), with \( L \ll n \), then the leading term of \( b^* \) drops further to \( n \times 4/(L - 1) \). The results for intermediate values of \( L \), with the possibility of \( L > n \), are presented in Supplementary Method 1.4.

There are also alternative intercommunity connection schemes that offer a marked reduction in \( b^* \). For example, if there is one broker node between the gate nodes and the broker is connected to \( m > 1 \) peripheral leaf nodes (Fig. 1d), then the leading term of \( b^* \) is given by \( n(m + 2)(m + 5)/(m(m + 3)) \), which is linear in \( n \).

In actual settings, often there are more than two communities (local social networks and production units, among others). Urbanization has led to a proliferation of diverse subcultures and enhanced interaction and diffusion between them as a daily principle of contemporary life. In organizational settings, ‘network brokers’ can bridge existing structural holes and connect multiple segregated sectors and facilitate cooperation among them. A simple example of such a setting would be a star of cliques: \( m > 2 \) communities connected via a highly central broker node (Fig. 1e). With this \( m \)-community structure, with \( m > 2 \) communities, \( b^* \) has the leading term \( n \times m(m - 2) \). Linear growth in community size \( n \) indicates a substantial improvement over a single community or two communities is attained.

Another interconnection scheme of multiple cohesive communities is the ‘caveman graph’ from the sociological literature (Fig. 1f). With \( L > 2 \) cliques, situated on a ring, the leading term of \( b^* \) is given by \( n \times L/(L - 2) \), which is linear in clique size.

Cliques can also be organized hierarchically, such as in modern organizational bureaucracies (Fig. 1g). In this topology, a head node is connected to \( q \) subsidiary nodes, and each of them is connected to \( q \) cliques of size \( n \). If \( q \ll n \), then the leading term of \( b^* \) grows linearly with clique size \( n \).

A star graph comprises a hub and \( n \) leaf nodes connected to the hub. In this strictly centralized system, natural selection does not promote cooperation regardless of \( b \). Similar to the case of cliques, stars can be connected to promote collective cooperation. If we have two stars, one with \( n \) leaf nodes and the other with \( m \) leaf nodes (Fig. 2a), then if we connect the hubs, \( b^* \) for large \( n \) approaches a constant \( (8 + \alpha + 1/\alpha)/4 \). The smallest possible \( b^* \) for two stars is 5/2, which pertains to \( \alpha = 1 \) (identical stars). The independence from network size is a remarkable feature that star structures exhibit.

If we connect the hubs via one intermediary broker node, we get \( b^* \approx (10 + \alpha + 1/\alpha)/4 \). For two identical stars, this simplifies to \( b^* \approx 3 \). We can also connect the hubs via a chain of \( L \) intermediary brokers, such as in a chain-of-command structure with a decision-making unit at the top and an implementation unit at the bottom. For \( L \geq 1 \), the leading term of \( b^* \) is given by \( (8 + 2L + \alpha + 1/\alpha)/(L + 3) \). In all these cases, it is remarkable that for a large network size, \( b^* \) tends to be a constant. This independence from network size evinces the high merit of locally star-like structures in the promotion of cooperation.

In many actual settings, star-like structures are not directly connected as we envisaged above. Rather, global hubs are connected to local large-scale hubs, which are in turn connected to local peripheral nodes. This leads to a hierarchical organization: the head unit connects to numerous subsidiary units, each of them connect in turn to subordinate units, and so on. To study this interconnection scheme, we consider graphs with mega-hubs and hubs in a nested manner. We consider only two levels, although the calculation can be, in principle, extended to more. Out of the \( n \) total leaf nodes of a star graph, we take \( n_g \) of them and attach \( n_d \) nodes to each (Fig. 2b), where the \( g \) subscript denotes ‘gate’ and the \( d \) subscript denotes ‘dangling.’ The total number of nodes will be \( 1 + n + n_g + n_d \), and the number of links is \( n + n_g \). The full expressions for \( b^* \) are long (see Supplementary Method 1.3), but simplifications can be obtained in some interesting limits. We consider the case in which the number of leaf nodes is much larger than the number of hubs. This is the case in many actual settings, to the extent that the marked imbalance between the latter two numbers constitutes the cornerstone of many egalitarian social discourses and movements. If we have \( n_g \ll n \), then the leading term of \( b^* \) approaches 3/2. Whether \( n_g \) and \( n \) are of the same order of magnitude or whether we have \( n \ll n_g \), only affects the second-leading order terms. This leading behaviour of \( b^* \) is particularly interesting because the average degree approaches two in these cases, and \( b^* \) being less than the average degree is a rare property of graphs. Hence, we can dub these structures ‘super-promoters’ of cooperation.

Fig. 2 | Super-promoters of cooperation. Star graphs represent extreme core-periphery structures where a central node is connected to many leaf nodes. Although a single star hinders cooperation, connecting stars promotes cooperation. All reported critical benefit-to-cost ratios, \( b^* \), pertain to the limit of a large population size. Exact formulas are shown in Supplementary Methods. a, Two stars, one with \( n \) leaf nodes and the other with \( m \) leaf nodes. b, An imperfect meta-star: a central node (pink) has \( n \) peripheral nodes, \( n_g \) of them are hubs (orange) and connected to \( n_d \) dangling nodes (cyan), while \( n - n_g \) of them are not connected to any subsequent nodes (blue). If \( n_g \ll n \), then \( b^* \) tends to be 3/2 and the average degree tends to be 2. Thus, the structure is a super-promoter of cooperation, as \( b^* \) is less than the average degree. c, The perfect meta-star is a hierarchical structure with a head node (blue) connected to \( n \) subsidiary nodes (\( n_s \)), each of them connected to \( n_s \) subsidiary nodes, each of them connected to \( n_d \) dangling nodes. d, A more flat hierarchical structure: there are \( m \) head nodes connected on a ring, each with \( n \) peripheral nodes. For \( m \ll n \), this graph becomes a super-promoter of cooperation, outperforming the strict hierarchy.
We can readily generalize these results to the fully hierarchical structure (where \( n_c = n \)), that is, a star of stars (Fig. 2c). A mega-hub is connected to \( n \) hubs, which are each connected to \( n_c \) leaf nodes. In the limit of \( n \ll n_c \), \( b^* \) approaches two, which indicates that this structure is a strong promoter of cooperation.

Hierarchies can also be more ‘flat’, which is getting popular in certain management approaches\(^{21,22}\). The simplest model would be to have the upper layer of nodes connect horizontally instead of hierarchically. We consider the simple case in which the hubs of \( m \) stars, each with \( n \) leaf nodes, are connected on a ring (Fig. 2d). For large \( n \), the leading term of \( b^* \) approaches \((3m-1)/(2m-2)\), which is independent of \( n \). This means that for large \( m \), the value of \( b^* \) approaches \(3/2\). The average degree in this limit approaches two. Hence, a ring of stars is another super-promoter of cooperation.

A rich-club network is one comprising a small, dense core of connection-rich high-degree nodes and a large, sparse periphery. These structures are found across social and technological networks. The notion of ‘oligarchy’ in institutions and organizations is usually linked to structures that can be characterized by such a rich-club feature\(^1\). Other examples with this feature include the social network of company executives and directors (within-company\(^23\), national inter-company\(^24\) and international inter-company\(^25\)), the collaboration network between academics\(^26\) and the Internet\(^27\).

As a simple example with this characteristic, we consider a clique of \( n_c \) nodes (where \( c \) denotes ‘core’) and \( n_p \) peripheral nodes. Each core node is connected to every other core node and every peripheral node. Each peripheral node is connected to every core node but to none of the other peripheral nodes. In the special case of \( n_c = 1 \), this becomes a star graph. For a single rich-club network, natural selection does not favor cooperation, regardless of \( b \). Similar to the case of a single clique, single rich-club networks promote spite.

To improve the situation, we connect two rich-club networks by connecting a ‘gate’ node in the first core to a gate node in the second core (Fig. 3a). An actual example of conjoining rich clubs via cores is that of director networks of different companies that are often connected, and they do so predominantly via their cores, rather than the peripheries—creating ‘interlocked directorates’\(^27\). In the simple case of two identical rich-club networks with \( n_c \ll n_p \) (a small core and a large periphery), the leading term of \( b^* \) is given by \( 4 n_c - 3/2 \), which is a linear function of \( n_c \). That is, the leading behaviour in the large-\( n_c \) limit only depends on the number of core nodes and is independent of the number of peripheral nodes. In the case of \( n_c = 1 \), this leading term is \( 5/2 \), which is consistent with our previous findings for star graphs. For \( n_c = 2 \), the leading term of \( b^* \) is \( 13/2 \). The results point out a remarkable feature of these structures: when the periphery is large, the fate of the collective outcome is determined solely by the core.

In a bipartite network, nodes can be divided into two distinct groups, where there is no intra-group link. For example, heterosexual marriage networks comprised two disjoint sets; males only connected to females and vice versa. Other examples include buyer–seller\(^28\) and employer–employee\(^29\) bipartite networks.

Here, we present the results for the simplest case of a bipartite graph, which is analytically tractable: we consider a complete bipartite graph. A complete bipartite network is one that has two groups, and each node is connected to every node in the other group but to no node in its own group. Natural selection does not promote cooperation on a complete bipartite graph, regardless of \( b \). However, if we connect two bipartite networks, the situation improves (Fig. 3b). Consider a bipartite graph comprising two groups of nodes with sizes \( n_x \) and \( n_y \), respectively. Suppose we connect two identical such bipartite graphs by connecting a type-x node in the first graph to a type-x node in the second graph. In the special case of \( n_x = n_y = n \), \( b^* \) grows linearly with \( n \). The leading term of \( b^* \) in this case is given by \( 2n \). Alternatively, if \( n_x \ll n_y \), then the leading term of \( b^* \) only depends on \( n_y \), and is given by \( 4n_y - 3/2 \). Hence, similar to rich clubs, if a large group of nodes are not interconnected within themselves and are all connected only to another small group of nodes, the collective outcome will be determined by that small group.

As actual social networks typically have more randomness than the ideal structures considered above, we investigate random networks to check whether they have qualitatively similar properties. Our first test (see Supplementary Method 1.9) is to add structural noise to the above-considered topologies and verify that the \( b^* \) values are indeed robust against structural deviations. For the next check, we investigate how conjoining cooperation-inhibiting random networks can promote cooperation. We generate 10 random Erdős–Rényi (ER) graphs\(^30\), with values of \( b^* \) that are undesirable for cooperation: negative (promoting spite) or highly positive (hindering cooperation). Network size is fixed at 40. There are 55 possible network pairs (45 pairs in which the two networks are different and 10 pairs in which they are identical), and there are 1,600 ways to conjoin two networks via one gate node in each. We calculate the median value of \( b^* \) among all these possible conjoinings for each pair of networks. The lower triangle in Fig. 4a presents the resulting \( b^* \) of the conjoined network against the \( b^* \) of the first and the second network. The upper triangle presents the results for the same procedure, except the gate nodes are connected via one broker node, instead of being directly connected. It can be seen that, in most cases, a substantial improvement is achieved in both conjoining schemes. The most resistant case is the one with \( b^* = -50 \). Note that networks whose \( b^* \) is negative are promoters...
Fig. 4 | Conjoining random graphs and empirical networks. a, Conjoining two ER random graphs directly (bottom triangle, orange values) and via one broker node (upper triangle, cyan values). The critical benefit-to-cost ratio, $b^*$, of each graph separately (as shown on the x and y axes) is either negative (promoting spite) or highly positive (promoting cooperation but at a very high benefit-to-cost ratio). The graph size is 40. In all cases, connecting two graphs leads to critical benefit-to-cost ratios of order $n$. Thus, we find that conjoining cooperation-inhibiting random graphs also promotes cooperation. b, Extension to more than two graphs: four ER graphs with link-formation probabilities of 0.8 (top left, promoting spite), 0.7 (top right, promoting spite), 0.45 (bottom right, promoting spite) and 0.35 (bottom left, promoting cooperation, with $b^* \approx 43$). The intercommunity link probability is 0.01. The very few intercommunity links engender a considerable improvement: the overall structure promotes cooperation $b^* \approx 14$, which is markedly better than each individual group. c, Conjoining two scale-free networks generated by the KE model. The orange values pertain to direct conjoining, and blue values pertain to conjoining via one broker node. d, An example network with the community structure generated by the LFR benchmark. Community structure is a ubiquitous feature of actual social networks; there are cohesive friendship groups connected by long ties. e–h, We use four empirical datasets pertaining to friendship networks. We used standard community detection algorithms to partition the dataset into communities and calculated the critical benefit-to-cost ratio for the whole network and for each community. In every case, the whole network was better than individual subnetworks in promoting cooperation.
of spite, and the closer to zero the value of \( b^* \) is, the more strongly the structure promotes spite. The results indicate that, if the spite-promotion capacity of either group is high, conjoining them would be less helpful collectively. In Supplementary Method 1.7, we present these results using \( 1/b^* \) instead of \( b^* \) to measure the conduciveness to cooperation. In Fig. 4b, we illustrate that the conjoining mechanism also works for more than two ER networks. In the example case shown, three of the four ER networks promote spite and one of them promotes cooperation with \( b^* \approx 43 \). Creating intercommunity links between these four groups with a probability of 0.01 begets a marked improvement: the overall structure has \( b^* \approx 14 \), which is considerably better than each of the individual groups for promoting cooperation. A generalization of this procedure gives rise to the stochastic block model, which we investigate in Supplementary Method 1.8.

The same conjoining procedure is applicable to networks with heavy-tailed degree distributions, which emulate actual social networks more realistically than ER networks. Here, we use the model proposed by Klemm and Eguíluz (KE)\(^\text{13} \) to generate scale-free networks with both small-world property and high clustering coefficient, which are both ubiquitous features in social networks. The results are presented in Fig. 4d. Conjoining every pair of networks produces a composite network with positive \( b^* \). In Supplementary Method 1.7, we present the results for four additional scale-free models. The results are qualitatively similar and the improvement in \( b^* \) via conjoining ensues consistently.

To study the effect of community structure on the cooperative outcome, we employed the Lancichinetti–Fortunato–Radicchi (LFR) benchmark\(^\text{14} \) that is used for comparing community-detection algorithms. The procedure generates networks with community structure in which the degree distribution within each community and the distribution of community sizes are both heavy tailed. Figure 4e depicts an example case with 100 nodes divided into three communities. The degree distribution is scale-free with an exponent of two. The community sizes are 10, 23 and 67. Only the largest community has a positive \( b^* \), with \( b^* \approx 99 \). The composite network (with a mixing parameter of 0.1) has \( b^* \approx 35 \). In Supplementary Method 1.8, we provide a systematic investigation for LFR networks and show that, consistent with the above findings, when communities are not conducive to cooperation, sparse interconnections tend to generate composite networks better than the individual modules.

We can apply the same mathematical formalism to real-world social network data. We use offline social networks that pertain to friendships, to ascertain that cooperative dynamics would be reasonable. We use two children friendship networks of fourth-grade and fifth-grade students\(^\text{33,34} \). In this dataset, for the third grade, no reasonable. We use two children friendship networks of fourth-grade and fifth-grade students\(^\text{33,34} \). The results are presented in Fig. 4f. Conjoining every pair of networks produces a composite network with positive \( b^* \). In Supplementary Method 1.7, we present the results for four additional scale-free models. The results are qualitatively similar and the improvement in \( b^* \) via conjoining ensues consistently.

Each population structure can be quantified according to its intrinsic propensity to promote cooperation (paying a cost to benefit others) or spite (paying a cost to harm others)\(^\text{3} \). Here, we report the observation that sparsely conjoining cooperation-inhibiting structures tend to produce cooperation-promoting structures. We have explored this effect when joining together fully connected cliques, star-like structures (which are dominated by a single individual), rich clubs and even random graphs. We have found the phenomenon in examples of real social networks that already consist of conjoined substructures.

In our findings, conjoining two graphs that are already favourable for cooperation always results in a cooperation-promoting composite structure, although sometimes the composite graph might not promote cooperation as strongly as the two individual graphs did. However, we did not find any example in which the composite graph would inhibit cooperation, that is, either with \( b^* \) significantly larger than those of the two initial graphs or with a negative \( b^* \). We investigated random and non-random graph families considered in this paper, and several others.

An extension to our work would be finding better conjoining schemes for cliques. Here, we showed that conjoining cliques in the manners described above results in composite networks that are considerably better than individual cliques. These conjoining methods yield \( b^* \) values that grow linearly with \( n \). For very large networks, this improvement might still not be enough. A valuable extension would be to find structures that, similar to the case of stars and rich clubs, would produce \( b^* \) that reaches a constant for large clique size.

We note that evolutionary graph theory, which we used in this paper, is a general approach to study the effect of population structure on natural selection. It is not limited to any particular game and not restricted to one-shot interactions. The results are generalizable to any matrix game (see Methods). Hence, the competing strategies could instantiate repeated interactions and conditional behaviour\(^\text{15} \). Extensions of evolutionary graph theory can be used to study direct reciprocity with crosstalk\(^\text{16} \) and indirect reciprocity with optional interactions and private information\(^\text{17} \). Conversely, there are social settings that our model is not applicable to. For example, if each individual interacts with only a subset of its neighbours, then exclusion and inclusion become essential elements of network power. This is an important feature in Network Exchange Theory\(^\text{41} \). In this case, broker nodes have leverage over others due to the high exclusion–inclusion asymmetry. Our model does not consider the possibilities of exclusion and inclusion, and each player plays with every neighbour. Thus, an interesting extension to the present paper would be to study analytically the aforementioned effects of exclusion and inclusion in a game-theoretical setting to build on the previous experimental work, particularly Network Exchange Theory.

Finally, we highlight that our results are qualitatively consistent with several simulation studies in the literature across different contexts: cooperation is promoted by interdependence between networks in spatial public goods games and the prisoner’s dilemma on interdependent networks\(^\text{42–44} \), even if it is endogenous and interpopulation links are only rewarded to high-payoff individuals\(^\text{45} \). The same is true if multiple types of interactions are considered, resulting in a multiplex network\(^\text{46} \).

Our findings suggest a framework for how to build societal structures that effectively promote cooperation, and, together with the ensemble of previous results in the literature, they engender hope regarding the increasing interconnection of the contemporary world.

**Methods**

We follow a recently discovered framework for unweighted, undirected graphs without self-loops. Let us denote the degree of node \( x \) with \( k_x \), and its set of neighbours by \( N_x \). Then, we define \( p_x \), as the probability that a random walk of length 2 initiated at node \( x \) will terminate at node \( x \):

\[
p_x = \frac{1}{k_x} \sum_{y \in N_x} \frac{1}{k_y}
\]
We then solve the following system of $\binom{n}{2}$ linear equations for symmetric quantities $r_{xy}$, which are the meeting times of two random walkers initiated at nodes $x$ and $y$:

$$r_{xy} = r_{yx} = (1 - \delta_{xy}) + \frac{1}{2k_{x}} \sum_{y \neq x} r_{xy} + \frac{1}{2k_{y}} \sum_{x \neq y} r_{yx}$$

(2)

Here, $\delta_{xy}$ equals unity if $x = y$ and is zero otherwise. Using these quantities, we define $r_{x}$ for each node as the expected meeting time of two random walkers initiated at node $x$ as follows:

$$r_{x} = \frac{1}{k_{x}} \sum_{y \neq x} r_{xy}$$

(3)

The necessary condition for cooperation to be favoured by natural selection is that $b(\sum p_{k} r_{k} - 2NK)$ is greater than $c(\sum r_{k} - 2NK)$. If the coefficient of $b$ in this inequality is non-positive, cooperation is never favoured. If the coefficient is positive, then the critical benefit-to-cost ratio is given by the following relation:

$$b^* = \frac{\sum r_{k} - 2NK}{\sum p_{k} r_{k} - 2NK}$$

(4)

The calculations for specific graphs discussed in the main text can be simplified utilizing their structural symmetry. For example, for a single community (a complete graph), there is only one variable: the meeting time between any pair of nodes (because $r_{xy}$ values are zero). For two communities connected directly by a link, there are only four distinct values for $r_{xy}$: the meeting time between two commuters, between a commuter and the gate node of the same community, between a commuter and the gate node of the other community and between the two gate nodes. This reduces equation (2) to a system of four equations with four unknowns.

The results are generalizable to arbitrary $2 \times 2$ games. For a game with strategies A and B and with a corresponding payoff matrix $R, S, T, P$, the condition that natural selection favours strategy A over strategy B in the limit of weak selection is: $(T - S) < (R - P)(b^* + 1)/(b^* - 1)$. For the KE networks used in Fig. 4c, we used the model of Klemm and Eguíluz. We generated many networks, with the crossover parameter $p$ and the number of initial active nodes $m$ both selected randomly in their valid ranges. We selected six networks whose $b^*$ differed from the corresponding values used in Fig. 4b by <5%.

**Reporting Summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

**Code availability.** For the LFR benchmark, we used the publicly available code that the authors of ref. 32 have provided: https://sites.google.com/site/lancichinetti31. For the coalescing random walks framework, the code for computing $b^*$ is publicly available in Zenodo at https://doi.org/10.5281/zenodo.376933.

**Data availability.** All the network datasets used in this paper are freely and publicly available in the Colorado Index of Complex Networks (ICON) collection: https://icon.colorado.edu

Received: 9 November 2017; Accepted: 22 May 2018; Published online: 2 July 2018
Acknowledgements
This work was supported by the James S. McDonnell Foundation (B.F.), NSF grant 1715315 (B.A.) and the John Templeton Foundation (M.A.N.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. B.F. thanks S. Rytina for insightful and stimulating conversations.

Author contributions
All authors contributed to all aspects of the paper.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41562-018-0368-6.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to B.F.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Corresponding author(s): Babak Fotouhi

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- n/a Confirmed
- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
- Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated
- Clearly defined error bars
- State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection | No data collection was done in this study.

Data analysis | Numerical solution to the linear equations were done in MATLAB_R2015b and Mathematica 10.4.1.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All the network data sets used in this paper are freely and publicly available in The Colorado Index of Complex Networks collection: https://icon.colorado.edu
**Field-specific reporting**

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- [ ] Life sciences
- [x] Behavioural & social sciences
- [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](http://nature.com/authors/policies/ReportingSummary-flat.pdf)

**Behavioural & social sciences study design**

All studies must disclose on these points even when the disclosure is negative.

| Study description                       | This study provides a mathematical (hence quantitative) model of cooperation on social networks. Most of our numerical analyses are on theoretical graphs and simulated networks. To show generalizability, we also considered two publicly-available network data sets. |
|-----------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Research sample                         | We used existing network data sets publicly available in The Colorado Index of Complex Networks collection: https://icon.colorado.edu                                                                 |
| Sampling strategy                       | No sampling was done in this study.                                                                                                                                                                  |
| Data collection                         | No data collection was performed in this study.                                                                                                                                                      |
| Timing                                  | No data collection was performed in this study.                                                                                                                                                      |
| Data exclusions                         | No data was excluded.                                                                                                                                                                               |
| Non-participation                       | The study did not involve participation of any human or non-human subjects.                                                                                                                           |
| Randomization                           | The study did not involve participation of any human or non-human subjects.                                                                                                                           |

**Reporting for specific materials, systems and methods**

| Materials & experimental systems | Methods |
|----------------------------------|---------|
| n/a                              | n/a     |
| ![Unique biological materials](on) | ![Involved in the study](on) |
| Antibodies                       | ![Involved in the study](on) |
| Eukaryotic cell lines            | ![Involved in the study](on) |
| Palaeontology                    | ![Involved in the study](on) |
| Animals and other organisms      | ![Involved in the study](on) |
| Human research participants      | ![Involved in the study](on) |
| ![ChIP-seq](on)                  | ![Involved in the study](on) |
| ![Flow cytometry](on)            | ![Involved in the study](on) |
| ![MRI-based neuroimaging](on)    | ![Involved in the study](on) |