

# Evolution of Cooperation on Large Networks with Community Structure

Babak Fotouhi<sup>1,2</sup>, Naghmeh Momeni<sup>1,3</sup>, Benjamin Allen<sup>1,4,5</sup>, Martin A. Nowak<sup>1,6,7</sup>

<sup>1</sup> *Program for Evolutionary Dynamics, Harvard University, Cambridge, MA, USA*

<sup>2</sup> *Institute for Quantitative Social Sciences, Harvard University, Cambridge, MA, USA*

<sup>3</sup> *Massachusetts Institute of Technology (MIT) - Sloan School of Management, Cambridge, MA, USA*

<sup>4</sup> *Department of Mathematics, Emmanuel College, Boston, MA, USA*

<sup>5</sup> *Center for Mathematical Sciences and Applications, Harvard University, Cambridge, MA, USA*

<sup>6</sup> *Department of Mathematics, Harvard University, Cambridge, MA, USA*

<sup>7</sup> *Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA*

Cooperation is a major factor in the evolution of human societies. The structure of social networks, which affects the dynamics of cooperation and other interpersonal phenomena, have common structural signatures. One of these signatures is the tendency to organize as groups. This tendency gives rise to networks with community structure, which are composed of distinct modules. In this paper, we study analytically the evolutionary game dynamics on large modular networks in the limit of weak selection. We obtain novel analytical conditions such that natural selection favors cooperation over defection. We calculate the transition point for each community to favor cooperation. We find that a critical inter-community link creation probability exists for given group density, such that the overall network supports cooperation even if individual communities inhibit it. As a byproduct, we present solutions for the critical benefit-to-cost ratio which perform with remarkable accuracy for diverse generative network models, including those with community structure and heavy-tailed degree distributions. We also demonstrate the generalizability of the results to arbitrary two-player games.

## INTRODUCTION

Cooperation is a central tenet of social life of many species. When individuals cooperate, they pay a cost to help one another. The dynamics of interpersonal cooperation is affected by the structure of social networks, which characterize patterns of interpersonal interaction and exchange [1–5]. The problem of cooperation is also widely studied in mathematical and physical sciences [6].

In this paper, we focus on a particular feature of social networks, namely, community structure. This structural feature is ubiquitous in social networks, and is associated with a tendency of societies to organize as modules, or communities, with disparate intergroup and intragroup connectivity [7]. Community structure is measured and studied in many species, such as dolphins [8], chimpanzees [9], macaques [10], sharks [11], various pollinator species [12], and even relatively solitary species such as desert tortoise [13]. Community structure regulates the interpersonal connections and thereby affects the dynamical processes taking place on social networks. For example, the presence of community structure affects the epidemic spreading of infectious disease for various animal species [14]. Networks with high modularity buffer the spread of ecological perturbation impact across the whole population [15]. Also in the social life of humans, community structure is shown to promote the spread of health behaviors [16] and the diffusion of information online [17]. The effect of community structure on cooperative dynamics has also been

studied under diverse simulation settings and modeling assumptions[18–21]. Here we aim to provide the first analytical results to study the effect of community structure on the evolution of cooperation. We first introduce a potent approximation framework which, as we shall demonstrate, performs with remarkable accuracy for a wide array of network models with diverse properties, including heavy-tailed degree distributions, small-world properties, varying levels of clustering, and community structure. We use this framework to put in crisp focus a standard generative network model, namely, the Stochastic Block Model (SBM). We provide an explicit solution for the conditions such that natural selection favors cooperation, and then generalize the results to any  $2 \times 2$  evolutionary game.

## MODEL OF SOCIAL INTERACTIONS

We use the framework of evolutionary graph theory [1]. We consider a generic two-player game with two available strategies, denoted by C and D. Both players receive  $R$  if they mutually choose C, and receive  $P$  if they mutually choose D. If one player chooses C and the other chooses D, the C-player receives  $S$  and the D-player receives  $T$ . The strategy of node  $x$  is denoted by  $s_x \in \{0, 1\}$ , where 1 corresponds to strategy C and 0 to strategy D. We denote the set of network neighbors of node  $x$  by  $\mathcal{N}_x$ . We consider averaged payoffs, where the

payoff of node  $x$  with degree  $k_x$  is given by

$$f_x = \frac{1}{k_x} \sum_{y \in \mathcal{N}_x} s_x [Rs_y + S(1 - s_y)] + (1 - s_x) [Ts_y + P(1 - s_y)]. \quad (1)$$

At each timestep, a random individual is chosen to update its strategy via copying that of a neighbor. The probability that neighbor  $x$  is copied is proportional to  $1 + \delta f_x$ , where  $f_x$  is the payoff of  $x$  and  $0 < \delta \ll 1$  models the selection strength, that is, higher  $\delta$  indicates stronger social learning via observing the payoff of peers. The case  $\delta = 0$  is equivalent to the *voter model* [22–24]. The weak-selection limit considered here can be viewed as the first-order correction to the voter model. We seek an analytical condition such that natural selection favors the fixation of C over fixation of D. According to the Structure Coefficient Theorem [25], this happens if  $(R - P)\sigma > (T - S)$ , where  $\sigma$  is the structure coefficient, which is independent of the game and only depends on the network structure. Thus, it suffices to calculate  $\sigma$  for the given network. This can be done *exactly* by considering a mathematical equivalence between the evolutionary game dynamics and that of coalescing random walks [26]. Below we highlight the high computational cost of this method for large networks, propose a mean-field approximation to ameliorate the situation for large networks, and use it to obtain remarkably accurate solutions for networks with community structure. We also explain the referent of the term ‘mean-field’ in the proposed solution.

## GENERAL FRAMEWORK

To obtain  $\sigma$  for generic  $2 \times 2$  games, we apply the methodology of [26] to unweighted undirected graphs. We analyze the ‘donation game’ version of the Prisoner’s Dilemma, with  $R = b - c$ ,  $S = -c$ ,  $T = b$ , and  $P = 0$  and later discuss how  $\sigma$  can be derived from the results. For notational brevity, we set  $c = 1$  without loss of generality. This choice is simply equivalent to a rescaling of every payoff value. We shall discuss how the results for generic  $c$  can be readily obtained from provided results.

The payoff of node  $x$  is:

$$f_x(t) = -s_x(t) + \frac{1}{k_x} \sum_{y \in \mathcal{N}_x} b s_y(t). \quad (2)$$

For the update of node  $x$ , we have:

$$\begin{aligned} \mathbb{E}[s_x(t+1)] &= (1 - \frac{1}{N})s_x(t) \\ &+ \frac{1}{N} \sum_{y \in \mathcal{N}_x} \frac{1 + \delta f_y(t)}{\sum_{z \in \mathcal{N}_x} [1 + \delta f_z(t)]} s_y(t). \end{aligned} \quad (3)$$

The first term on the right hand side corresponds to the event that node  $x$  is not chosen to update its strategy. The second term corresponds to the expected value of  $s_x$  conditional upon node  $x$  having been chosen to update. In the limit of weak selection, expanding to the first order of  $\delta$  after multiplying both sides by  $k_x$ , we get:

$$\begin{aligned} \mathbb{E}[k_x s_x(t+1)] &= \frac{1}{N} \left[ \sum_{y \in \mathcal{N}_x} s_y(t) + \delta \sum_{y \in \mathcal{N}_x} f_y(t) s_y(t) \right. \\ &\quad \left. - \sum_{y, z \in \mathcal{N}_x} \delta \frac{s_y(t) f_z(t)}{k_x} \right] + O(\delta^2) + (1 - \frac{1}{N})k_x s_x(t). \end{aligned} \quad (4)$$

Now we define  $\psi(t) := \sum_x k_x s_x(t)$ . Summing (4) over all nodes, we see that in the zeroth-order dynamics of the system (i.e.,  $\delta = 0$ , corresponding to the voter model), the expected value of  $\psi(t)$  is a conserved quantity [22, 24, 27]. The fixation probability can be obtained by equating the average value of  $\psi$  over every initial mutant placement with its expected value as  $t \rightarrow \infty$ . The expected first-order change of  $\psi$  is

$$\begin{aligned} \Delta\psi^{(1)}(\vec{s}) &= \frac{\delta}{N} \left[ \sum_x \sum_{y \in \mathcal{N}_x} f_y s_y - \sum_x \sum_{y, z \in \mathcal{N}_x} \frac{s_y f_z}{k_x} \right] \\ &= \frac{\delta}{N} \left[ \sum_x k_x f_x s_x - \sum_x \sum_{y, z \in \mathcal{N}_x} \frac{s_y f_z}{k_x} \right] \\ &= \frac{\delta}{N} \sum_x \left[ k_x s_x \left( -s_x + \frac{b}{k_x} \sum_{y \in \mathcal{N}_x} s_y \right) \right. \\ &\quad \left. - \sum_{y, z \in \mathcal{N}_x} \frac{s_y \left( -s_z + \frac{b}{k_z} \sum_{w \in \mathcal{N}_z} s_w \right)}{k_x} \right] \end{aligned} \quad (5)$$

Exchanging the summation order, denoting the first moment of the degree distribution (the average degree) by  $\mu_1$ , we get

$$\begin{aligned} \Delta\psi^{(1)}(\vec{s}) &= \frac{\delta}{N} \left[ - \sum_y k_y s_y^2 + b \sum_{x \in \mathcal{N}_y} s_y s_x \right. \\ &\quad \left. + \sum_{x \in \mathcal{N}_y} \sum_{z \in \mathcal{N}_x} \frac{s_y s_z}{k_x} - b \sum_{x \in \mathcal{N}_y} \sum_{z \in \mathcal{N}_x} \sum_{w \in \mathcal{N}_z} \frac{s_y s_w}{k_x k_z} \right]. \end{aligned} \quad (6)$$

We need to sum up these expected increments from  $t = 0$  up to  $t = \infty$ . Let  $\xi_x$  denote this expected total change for given initial condition in which only node  $x$  is C and all other nodes are D. The fixation probability for a given initial condition will then be  $(k_x + \xi_x)/(N\mu_1)$ . Averaging over all nodes, this becomes:

$$\rho_C = \frac{1}{N} + \frac{1}{N} \sum_x \xi_x. \quad (7)$$

The sum on the right hand side of (7) requires the calculation of temporal sum of the spin products on the right hand side of (6). Note that these summations are to be performed in the voter-model regime. That is, due to the factor  $\delta$ , we should only keep the summations in zeroth order. In [26, 28], it is shown that for any two nodes  $i$  and  $j$ , if we find the expected temporal sum of  $1/N - s_i s_j$  from  $t = 0$  to  $t = \infty$  under the voter-model dynamics and then average this sum over all single-node initial placements, the result is equal to  $\tau_{ij}/(2N)$ , where  $\tau_{ij}$  is the expected meeting time of two random walkers initiated at nodes  $i$  and  $j$ . These meeting times follow the following recurrence relation:

$$\tau_{ij} = \tau_{ji} = (1 - \delta_{ij}) \left[ 1 + \frac{1}{2k_i} \sum_{\ell \in \mathcal{N}_i} \tau_{\ell j} + \frac{1}{2k_j} \sum_{\ell \in \mathcal{N}_j} \tau_{\ell i} \right]. \quad (8)$$

Thus the random walk equivalence relates the fixation probability to the expected value of the meeting times of two random walkers initiated one, two, and three steps away on the network, corresponding to the three last terms on the right hand side of (6), respectively. Note that the expression ‘ $\ell$  steps away’ here refers to random-walk steps, rather than graph distance. So for example, node  $y$  is  $\ell$  steps away from node  $x$  if  $A_{xy}^\ell > 0$ , where  $A^\ell$  is the  $\ell$ -th power of the adjacency matrix. Using the meeting times which are the solutions to the system of equations (8), we define the quantity  $\tau_x$  as the expected remeeting time of two random walkers both initiated at node  $x$ :

$$\tau_x = 1 + \frac{1}{k_x} \sum_{y \in \mathcal{N}_x} \tau_{yx}. \quad (9)$$

We also define  $p_x = \sum_{y \in \mathcal{N}_x} 1/(k_x k_y)$ . Using these definitions, and after some algebraic simplifications, the fixation probability can be expressed in the following form:

$$\begin{aligned} \rho_C &= \frac{1}{N} + \frac{\delta}{2N} \left[ b \left( \sum_x \frac{k_x}{N\mu_1} \tau_x - 2 \right) \right. \\ &\quad \left. - \left( \sum_x \frac{k_x}{N\mu_1} \tau_x p_x - 2 \right) \right] + O(\delta^2). \end{aligned} \quad (10)$$

If we repeat the analysis for the fixation of D, we get the same result as (10), with the sign of the first-order term flipped. Setting  $\rho_C > \rho_D$ , we arrive at the condition such that natural selection favors the fixation of C over D. This outcome occurs if we have  $b > b^*$ , where  $b^*$  is the critical benefit-to-cost ratio and is given by:

$$b^* = \frac{\sum_x \tau_x k_x - 2N\mu_1}{\sum_x \tau_x k_x p_x - 2N\mu_1}. \quad (11)$$

The result for generic  $c$  is readily obtained by replacing  $b$  in the above equation by  $(b/c)$ . For any value of  $b/c$  below this value, natural selection favors defection over cooperation. Moreover, if the value of  $(b/c)^*$  for a given network turns out to be negative, this means that natural selection favors defection over cooperation regardless of  $b$  and  $c$ . This network version of the tragedy of the commons would be unsalvageable via modification of the payoff structure. Below we discuss an example in which cooperation can be rescued by structural intervention instead, that is, by bridging individually-uncooperative networks together.

## LARGE NETWORKS: MEAN-FIELD APPROXIMATION

The drawback of this exact framework is that solving (8) requires solving a system of  $N(N - 1)/2$  linear equations, which can be infeasible for large networks. The conventional Cholesky decomposition methods to solve this system have the complexity of order  $N^6$ , and faster techniques often require sparsity. So for large networks in general, obtaining the exact solution is computationally infeasible. To obviate this limitation, here we seek a mean-field approximation to obtain analytical results that are computationally feasible for large networks. We use the fact that, combining (9) and (8), the remeeting times of these random walkers satisfy the following equation:

$$\sum_x k_x^2 \tau_x = N^2 \mu_1^2. \quad (12)$$

Using this equation and assuming a mean-field approximation in which every  $\tau_x$  value is replaced by the average value over all  $x$ , we get:

$$\tau_x = N \frac{\mu_1^2}{\mu_2}, \quad (13)$$

where  $\mu_2$  is the second moment of the degree distribution. This substitution of the  $\tau_x$  values with their mean is the rationale behind the term ‘mean-field’ we employed, which is conventional in the mathematical literature of evolutionary games [2]. For the mean-field approximation of the fixation probability, we plug this into (10) and obtain:

$$\rho \approx \frac{1}{N} + \frac{\delta}{2N} \left[ b \left( N \frac{\mu_1^2}{\mu_2} - 2 \right) - \left( \frac{\mu_1}{\mu_2} \sum_x k_x p_x - 2 \right) \right] \quad (14)$$

Noting that  $\sum_x k_x p_x = \sum_x \sum_{y \in \mathcal{N}_x} 1/k_y$  is equal to  $N$  for any network, we get:

$$\rho \approx \frac{1}{N} + \frac{\delta}{2N} \left[ b \left( N \frac{\mu_1^2}{\mu_2} - 2 \right) - \left( \frac{N\mu_1}{\mu_2} - 2 \right) \right] \quad (15)$$

Plugging this result into (11), we arrive at the critical benefit-to-cost ratio above which natural selection favors cooperation over defection:

$$b^* \approx \frac{N - 2 \frac{\mu_2}{\mu_1^2}}{\frac{N}{\mu_1} - 2 \frac{\mu_2}{\mu_1^2}} \quad (16)$$

Note that the previous approximations in the literature can be obtained as limiting cases. If the variance of the degree distribution is small, we can expand Equation (16) in terms of the degree variance:

$$b^* \approx \frac{N - 2}{N/\mu_1 - 2} + \frac{2(1 - \frac{1}{\mu_1}) \text{var}(k)}{\left(1 - 2\frac{\mu_1}{N}\right)^2} \frac{\text{var}(k)}{N} + O\left(\frac{\text{var}(k)}{N}\right)^2. \quad (17)$$

If we assume that the graph is regular with degree  $\mu_1$  (that is, all nodes have degree  $\mu_1$ ), then the first term on the right hand side of (17) is equivalent to existing results in the biological literature for the special case of bi-transitive graphs [29]. This also indicates that the error of using the result of Ref. [29] for heterogeneous (i.e., non-regular) graphs increases when network density is close to 1/2. If in addition to small variance, we take the limit of large  $N$ , the right hand side of Equation (16) tends to  $\mu_1$ . Thus we recover the well-known  $b/c > k$  rule in the literature for large regular graphs [30].

## NETWORKS WITH COMMUNITY STRUCTURE

A common analytical model to generate networks with community structure and to perform community detection is the Stochastic Block Model (SBM) [31–33]. Here we study analytically the evolution of cooperation on SBM networks, and then extend the results to arbitrary symmetric  $2 \times 2$  games.

We consider a stochastic block model [31, 32] with  $m$  equi-probable groups, intra-community link probability  $p$  and inter-community link probability  $q$ . In this model, for each node, the probability of being assigned to either of the  $m$  communities is assumed to be the same. For node  $x$  with degree  $k_x$ , denote the number of within-community neighbors by  $k_x^{\text{intra}}$  and denote the number of its neighbors in other communities by  $k_x^{\text{inter}}$ . For large  $N$ , the average degree  $\mu_1$ , which is the expected value of  $k_x$ , is

$$\mu_1 = \mathbb{E}(k_x^{\text{intra}}) + \mathbb{E}(k_x^{\text{inter}}) = p \frac{N-1}{m} + q(N-1) \left(1 - \frac{1}{m}\right). \quad (18)$$

Also because the inter and intra-community degree distributions are independent, the variance of  $k$  is the sum

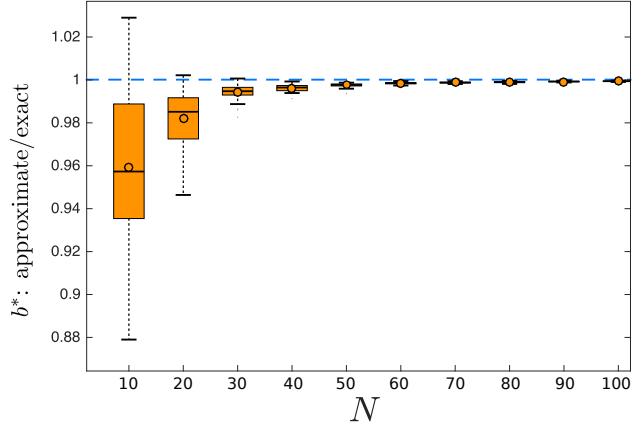


FIG. 1. Accuracy of the proposed mean-field approximation for  $b^*$  as a function of  $N$ . The network generation parameters are set to the example values of  $p = 0.7$ ,  $q = 0.1$ , and  $m = 3$ . The performance is consistently well for every parameter configuration tried, which will be demonstrated in Figure 3.

of the variance of  $k_x^{\text{intra}}$  and the variance of  $k_x^{\text{inter}}$ :

$$\text{var}(k) = p(1-p) \frac{N-1}{m} + q(1-q)(N-1) \left(1 - \frac{1}{m}\right). \quad (19)$$

Note that in SBM networks, although the assignment probabilities are uniform, this does not indicate homogeneous community size. The expected value of community sizes are the same, but at each realization of SBM networks, the community sizes are not necessarily the same. The probability of having  $m$  communities with identical size tends to zero as  $N$  increases for any  $m > 1$ . Combining (18) and (19), we obtain  $\mu_2$ , which we can insert into (15) to obtain the fixation probability. For the critical benefit-to-cost ratio, we insert the expression for  $\mu_2$  and  $\mu_1$  into (16). Defining  $\alpha := 1/m$  and  $\beta := 1 - 1/m$  for brevity, and after algebraic simplifications, we obtain:

$$b_{\text{SBM}}^* \approx \frac{N-2 - \frac{2}{N-1} \frac{\alpha p(1-p) + \beta q(1-q)}{(\alpha p + q\beta)^2}}{\frac{N/(N-1)}{\alpha p + \beta q} - 2 - \frac{2}{N-1} \frac{\alpha p(1-p) + \beta q(1-q)}{(\alpha p + \beta q)^2}}. \quad (20)$$

Figure 1 demonstrates that the approximation (20) has relative error less than 1% for network size as small as 40. In these network sizes, the exact method can be employed in reasonable time and thereby we have a benchmark to assess the solutions. Figure 1 demonstrates that for large networks, the error rate is remarkably small. So for large networks where the exact method become prohibitively costly, the mean-field approximation can be used to produce accurate results.

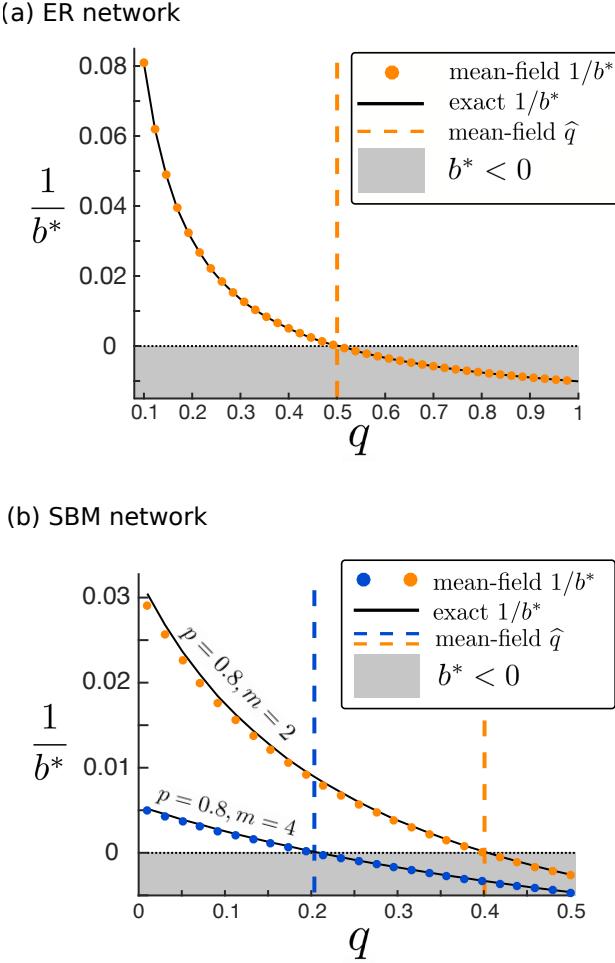


FIG. 2. (a) Accuracy of the proposed mean-field approximation for  $b^*$  for ER networks as a function of link creation probability  $p$ . As predicted by (21), there is a phase transition at  $\hat{p} = 1/2$ . That is, above  $\hat{p} = 1/2$ , there is a different cooperative phase: natural selection does not favor cooperation over defection *regardless* of the benefit-to-cost ratio. The mean-field prediction is in agreement with the exact results. (b) Accuracy of the proposed mean-field approximation for  $b^*$ . The network size is 100. The dashed lines depict the predicted threshold value  $\hat{q}$  as given by (22), agrees with the exact results for example parameter settings. For  $q > \hat{q}$ , natural selection promotes fixation of defection over cooperation *regardless* of the benefit-to-cost ratio. The intercepts match the result for  $\lim q \rightarrow 0^+$  obtained in (23).

The special case of  $p = q$  is equivalent to the Erdős-Rényi (ER) model [34]. Equivalently, we can set  $m = 1$ . In this special case, Equation (20) simplifies to:

$$b_{\text{ER}}^* \approx \frac{p(N^2 - 3N + 4) - 2}{(N - 2)(1 - 2p)}. \quad (21)$$

Figure 2a illustrates the accuracy of the proposed mean-field approximation (21) for ER networks. We plot  $1/b^*$  instead of  $b^*$  for better visualization. Equation (21) indicates a phase transition at  $\hat{p} = 1/2$ ; visible in Figure 2a. That is, in the ER model, the expected critical

benefit-to-cost ratio becomes negative if  $p > \hat{p}$ . In this regime, natural selection favors fixation of D over C for *any* benefit-to-cost ratio.

For the general case of  $p \neq q$ , too, we can find such a point of transition. We can find  $\hat{q}$  (the critical inter-community link probability above which natural selection favors fixation of D over C for any benefit-to-cost ratio) by setting the denominator of (20) equal to zero and solving the resulting quadratic equation for  $q$ . Expanding the solution for large  $N$ , we get:

$$\hat{q} = \frac{m - 2p}{2(m - 1)} + 2m \frac{\left(\frac{1}{2} - p\right)^2}{(m - 1)^2} \frac{1}{N} + O\left(\frac{1}{N^2}\right). \quad (22)$$

This result has an important consequence. Each of the  $m$  communities with intra-community link probability  $p$ , considered separately, is an ER network. So the expected value of the critical benefit-to-cost ratio for each of them is given by (21). Suppose the communities have  $p > 1/2$ , which, as discussed above, means that cooperation is not favored by natural selection for each individual community considered separately. We can then interconnect these communities under the SBM setting, with inter-community probability  $q$ . For  $q < \hat{q}$ , the critical benefit-to-cost ratio of the overall network is positive, despite individual communities inhibiting the fixation of cooperation. This confirms analytically the numerical observations about conjoining random networks [35].

In Figure 2b, we present the comparison of (20) with the exact results. The orange markers pertain to the example case of  $m = 4$  and  $p = 0.8$ , for which (22) gives  $\hat{q} \approx 0.4$ . The blue markers pertain to  $m = 2$  and  $p = 0.8$ , with  $\hat{q} \approx 0.2$ . The network size is 100 in both cases. The approximations are remarkably close to the exact values.

Of particular relevance for actual scenarios is the case where  $q \ll 1$ , which means that the communities are sparsely interconnected. In this regime, we can expand  $b^*$  as follows:

$$b^* = \frac{N(N - 2)p - 2m(1 - p)}{N(m - 2p) - 2m(1 - p)} + O(q). \quad (23)$$

The interesting result here is the existence of the zeroth-order term. This fact can be seen in Figure 2b as well. For  $N = 100, p = 0.8$  with  $m = 2$ , from (23) we get  $b^* \approx 200$ , and with  $m = 4$  we get  $b^* \approx 32.9$ . The inverse of these values are 0.0050 and 0.030, respectively, which correctly match the intercepts observed in Figure 2b. This confirms that in the limit as  $q \rightarrow 0^+$ ,  $b^*$  tends to a positive number. Thus, sparse interconnection of dense communities rescues cooperation. Note that this happens as long as the whole network is connected, so that  $b^*$  is well-defined. If  $q$  is exactly zero, the network will be segregated into disjoint components separated off, where  $b^*$  cannot be defined. So  $q$  can be as small as to ensure graph connectedness, but no

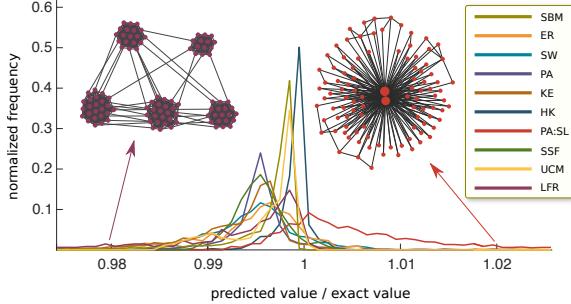


FIG. 3. (Color Online) The distribution of the ratio of the proposed mean-field approximation to the true value of  $b^*$ . The true value is obtained by solving the  $N(N - 1)/2$  linear equations as given by (8), which is infeasible for very large networks, but feasible for the size ranges of the test networks. The 10 network generation families are discussed in the text. Two example graphs with high heterogeneity are depicted to highlight the robustness of the proposed mean-field approximation to structural heterogeneity. By ‘heterogeneity’ we mean that the network properties of the nodes are different, as apposed to a regular network, such as a ring, in which every node exhibit similar structural properties.

smaller. The minimum value of  $q$  such that the whole network is connected asymptotically approaches zero as  $N$  tends to infinity. These results are conceptually consistent with the findings in the social sciences which underscore the important role of structural holes and boundary spanners [35, 36]. Nodes that function as network brokers connect the segregated patches of the network together and thereby facilitate collective cooperation. Broker nodes that were previously trapped in closed circles surrounded by defectors can now connect to nodes outside this circle, allowing cooperation to flow between clusters.

## EXTENSION OF RESULTS

**Extension to any game.** Returning to arbitrary  $2 \times 2$  games, according to the Structure Coefficient Theorem, the structure coefficient is defined in terms of  $b^*$  as follows:

$$\sigma = \frac{b^* + 1}{b^* - 1}. \quad (24)$$

For the special case of SBM or ER networks,  $\sigma$  can be obtained readily by inserting the corresponding values of  $b^*$  from Equations (20) and (21) into (24), respectively. For general networks, we use the mean-field value for  $b^*$  obtained in (16). The result simplifies to:

$$\sigma \approx \frac{N(\mu_1 + 1) - 4\mu_2/\mu_1}{N(\mu_1 - 1)}. \quad (25)$$

Natural selection favors the fixation of C over the fixation of D if  $(R - P)\sigma > (T - S)$ . Thus we have obtained analytical conditions such that natural selection favors

the fixation of one strategy over the other for any  $2 \times 2$  game.

**Extension to other network families.** We conclude by highlighting the accuracy of the proposed mean-field approximation for  $b^*$  via simulations and numerical results for other network families besides SBM and ER. In Figure 3, we plot the histogram of the ratio of the approximate  $b^*$  to the exact value (computed via solving the system of  $N(N - 1)/2$  linear equations as given by (8)) for 10 different network families. Besides SBM and ER, we consider the Small-world model [37], and 7 other network generation families, which all generate networks with heavy-tailed degree distribution. We use these additional 7 families as pessimistic scenarios regarding our mean-field approximation, because these models generate highly-heterogeneous networks. For the SBM model, we chose  $m$  uniformly from  $\{2, 3, 4, 5\}$ ,  $p$  uniformly in  $[0.1, 1]$ , and  $q$  uniformly in  $[.01, p]$ . For the small-world model [37], we chose the initial lattice degree uniformly from  $\times\{4, 8, 12\}$  and the link creation probability uniformly from  $[0, 0.1]$ . For ER networks, we chose the link formation probability uniformly in  $[0.2, 1]$ . For preferential attachment (PA) with linear kernel [38, 39], we randomly generate  $m$  between 1 and 5, and the kernel bias is generated randomly between 0 and 5. For the scale-free model of Holme and Kim (HK) [40], we chose the triad formation probability of the model uniformly in  $[0, 1]$ . For Klemm-Eguiluz (KE) scale-free model [41], we choose the cross-over probability parameter of the model uniformly in  $[0, 1]$ . For both models, we choose the number of initial connections of incoming nodes uniformly between 1 and 5. For the spatial scale-free model of Barthelemy [42] (SSL), we generated networks on a 2D lattice with distance decay parameter  $r_c$  chosen uniformly in  $[0, 0.2]$ . For the uncorrelated configuration model (UCM) [43], we chose the minimum number of connections uniformly between 1 and 5, and the exponent in the power-law degree distribution is chosen uniformly in  $[1, 4]$ . For super-linear preferential attachment [39] (PA:SL), where the attachment kernel depends on degrees as  $k^\theta$ , we chose the number of initial connections of incoming nodes uniformly between 1 and 4, and the exponent of the kernel uniformly in  $[0, 3]$ . Greater exponents produce networks with higher degree inequality. We also test the results on the LFR model [44], which is another conventional generative model for networks with community structure, besides SBM. Mathematical and statistical studies of networks involving community structure, such as those of inference and sampling, predominantly use the SBM model and its variants for analytical treatment. The usage of LFR models are comparatively more inclined towards simulations and numerical experiments. So in this section, we have also included LFR benchmarks in this paper for the sake of completeness and to confirm the reasonable performance of the

proposed solution on alternative network families with community structure. The are generated with the mixing parameter chosen uniformly at random in  $[0, 1]$ . The maximum degree  $k_{\max}$  was chosen uniformly between 1 and  $N - 1$ , the average degree was chosen uniformly between 1 and  $k_{\max}$ , the degree exponent was uniformly in  $[0, 3]$ , the exponent for the community size distribution was chosen uniformly in  $[0, 2]$ . We required that no community has less than 5% of the nodes, which also automatically introduces the cap  $N[1 - 0.05(m - 1)]$  on the maximum community size. For each family, we generate 10000 networks. For every generated network, the size is randomly chosen between 100 and 500, which is reasonably small so that the exact results could be feasibly calculated via solving (8). The histograms are highly concentrated around unity, which confirms the accuracy of the proposed mean-field approximations. Interestingly, the proposed approximation works well for networks with high structural heterogeneity, including those with heavy-tailed degree distributions.

## DISCUSSION

We presented accurate mean-field solutions for  $2 \times 2$  games on heterogeneous networks that determine which strategy is favored by natural selection. We have considered the limit of weak selection, which is amenable to analytical treatment. For general selection strength the problem is shown to be NP-hard [45]. We have utilized our solution to study the case of modular networks in detail, and have uncovered a network-structural phase transition which pertains to regions in which one of the strategies will not be favored by natural selection regardless of the payoff parameters. We have obtained similar results for ER networks as a special case. We have obtained analytical expressions for the inter-connection of segregated communities, which individually inhibit cooperation, but after interconnection, can collectively favor cooperation. This result in agreement with the previous results in the literature [18, 20], which obtain qualitatively similar results via various updating schemes and simulation parameters, and even those that consider many-player games [19]. Our analytical findings together with the previous simulation studies consistently highlight the cooperative advantage of sparsely interconnecting dense communities, and indicate that this advantage is a robust feature of cooperative dynamics on networks, which has notable real-world consequences.

**Acknowledgments.** This work was supported by a grant from the James S. McDonnell Foundation (N.M.), NSF grant 1715315 (B.A.), and the John Templeton Foundation (M.A.N.).

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

**Competing interests.** The authors declare no competing interests.

---

- [1] E. Lieberman, C. Hauert, and M. A. Nowak, “Evolutionary dynamics on graphs,” *Nature*, vol. 433, no. 7023, pp. 312–316, 2005.
- [2] G. Szabó and G. Fath, “Evolutionary games on graphs,” *Physics Reports*, vol. 446, no. 4–6, pp. 97–216, 2007.
- [3] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L. M. Floría, and Y. Moreno, “Evolutionary dynamics of group interactions on structured populations: a review,” *Journal of the Royal Society Interface*, vol. 10, no. 80, p. 20120997, 2013.
- [4] M. Perc and A. Szolnoki, “Coevolutionary games-a mini review,” *BioSystems*, vol. 99, no. 2, pp. 109–125, 2010.
- [5] B. Simpson and R. Willer, “Beyond altruism: Sociological foundations of cooperation and prosocial behavior,” *Annual Review of Sociology*, vol. 41, pp. 43–63, 2015.
- [6] M. Perc, J. J. Jordan, D. G. Rand, Z. Wang, S. Boccaletti, and A. Szolnoki, “Statistical physics of human cooperation,” *Physics Reports*, 2017.
- [7] M. Girvan and M. E. Newman, “Community structure in social and biological networks,” *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [8] D. Lusseau, B. Wilson, P. S. Hammond, K. Grellier, J. W. Durban, K. M. Parsons, T. R. Barton, and P. M. Thompson, “Quantifying the influence of sociality on population structure in bottlenose dolphins,” *Journal of Animal Ecology*, vol. 75, no. 1, pp. 14–24, 2006.
- [9] P. Kanngiesser, C. Sueur, K. Riedl, J. Grossmann, and J. Call, “Grooming network cohesion and the role of individuals in a captive chimpanzee group,” *American Journal of Primatology*, vol. 73, no. 8, pp. 758–767, 2011.
- [10] C. Sueur, O. Petit, A. De Marco, A. Jacobs, K. Watanabe, and B. Thierry, “A comparative network analysis of social style in macaques,” *Animal Behaviour*, vol. 82, no. 4, pp. 845–852, 2011.
- [11] J. Mourier, J. Vercelloni, and S. Planes, “Evidence of social communities in a spatially structured network of a free-ranging shark species,” *Animal Behaviour*, vol. 83, no. 2, pp. 389–401, 2012.
- [12] J. M. Olesen, J. Bascompte, Y. L. Dupont, and P. Jordano, “The modularity of pollination networks,” *Proceedings of the National Academy of Sciences*, vol. 104, no. 50, pp. 19891–19896, 2007.
- [13] P. Sah, K. E. Nussear, T. C. Esque, C. M. Aiello, P. J. Hudson, and S. Bansal, “Inferring social structure and its drivers from refuge use in the desert tortoise, a relatively solitary species,” *Behavioral Ecology and Sociobiology*, vol. 70, no. 8, pp. 1277–1289, 2016.
- [14] P. Sah, S. T. Leu, P. C. Cross, P. J. Hudson, and

S. Bansal, "Unraveling the disease consequences and mechanisms of modular structure in animal social networks," *Proceedings of the National Academy of Sciences*, p. 201613616, 2017.

[15] L. J. Gilarranz, B. Rayfield, G. Liñán-Cembrano, J. Bascompte, and A. Gonzalez, "Effects of network modularity on the spread of perturbation impact in experimental metapopulations," *Science*, vol. 357, no. 6347, pp. 199–201, 2017.

[16] D. Centola, "The spread of behavior in an online social network experiment," *Science*, vol. 329, no. 5996, pp. 1194–1197, 2010.

[17] L. Weng, F. Menczer, and Y.-Y. Ahn, "Virality prediction and community structure in social networks," *Scientific Reports*, vol. 3, p. 2522, 2013.

[18] Z. Wang, A. Szolnoki, and M. Perc, "Optimal interdependence between networks for the evolution of cooperation," *Scientific Reports*, vol. 3, p. 2470, 2013.

[19] Z. Wang, A. Szolnoki, and M. Perc, "Interdependent network reciprocity in evolutionary games," *Scientific Reports*, vol. 3, p. 1183, 2013.

[20] L.-L. Jiang and M. Perc, "Spreading of cooperative behaviour across interdependent groups," *Scientific Reports*, vol. 3, p. 2483, 2013.

[21] Z. Wang, A. Szolnoki, and M. Perc, "Rewarding evolutionary fitness with links between populations promotes cooperation," *Journal of Theoretical Biology*, vol. 349, pp. 50–56, 2014.

[22] V. Sood and S. Redner, "Voter model on heterogeneous graphs," *Physical Review Letters*, vol. 94, no. 17, p. 178701, 2005.

[23] T. Antal, S. Redner, and V. Sood, "Evolutionary dynamics on degree-heterogeneous graphs," *Physical Review Letters*, vol. 96, no. 18, p. 188104, 2006.

[24] V. Sood, T. Antal, and S. Redner, "Voter models on heterogeneous networks," *Physical Review E*, vol. 77, no. 4, p. 041121, 2008.

[25] C. E. Tarnita, H. Ohtsuki, T. Antal, F. Fu, and M. A. Nowak, "Strategy selection in structured populations," *Journal of Theoretical Biology*, vol. 259, no. 3, pp. 570–581, 2009.

[26] B. Allen, G. Lippner, Y.-T. Chen, B. Fotouhi, N. Momeni, S.-T. Yau, and M. A. Nowak, "Evolutionary dynamics on any population structure," *Nature*, vol. 544, no. 7649, pp. 227–230, 2017.

[27] B. Fotouhi and M. G. Rabbat, "Voter model with arbitrary degree dependence: clout, confidence and irreversibility," *The European Physical Journal B*, vol. 87, no. 3, p. 55, 2014.

[28] Y.-T. Chen *et al.*, "Sharp benefit-to-cost rules for the evolution of cooperation on regular graphs," *The Annals of Applied Probability*, vol. 23, no. 2, pp. 637–664, 2013.

[29] P. D. Taylor, T. Day, and G. Wild, "Evolution of cooperation in a finite homogeneous graph," *Nature*, vol. 447, no. 7143, pp. 469–472, 2007.

[30] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, "A simple rule for the evolution of cooperation on graphs and social networks," *Nature*, vol. 441, no. 7092, pp. 502–505, 2006.

[31] P. W. Holland, K. B. Laskey, and S. Leinhardt, "Stochastic blockmodels: First steps," *Social Networks*, vol. 5, no. 2, pp. 109–137, 1983.

[32] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, "Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications," *Physical Review E*, vol. 84, no. 6, p. 066106, 2011.

[33] B. Karrer and M. E. Newman, "Stochastic blockmodels and community structure in networks," *Physical Review E*, vol. 83, no. 1, p. 016107, 2011.

[34] P. Erdős and A. Rényi, "On random graphs," *Publicationes Mathematicae*, vol. 6, pp. 290–297, 1959.

[35] B. Fotouhi, N. Momeni, B. Allen, and M. A. Nowak, "Conjoining uncooperative societies facilities evolution of cooperation," *Nature Human Behavior*, vol. 2, pp. 492–499, 2018.

[36] J. C. Long, F. C. Cunningham, and J. Braithwaite, "Bridges, brokers and boundary spanners in collaborative networks: a systematic review," *BMC Health Services Research*, vol. 13, no. 1, p. 158, 2013.

[37] M. E. Newman and D. J. Watts, "Renormalization group analysis of the small-world network model," *Physics Letters A*, vol. 263, no. 4, pp. 341–346, 1999.

[38] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, "Structure of growing networks with preferential linking," *Physical Review Letters*, vol. 85, no. 21, p. 4633, 2000.

[39] P. L. Krapivsky, S. Redner, and F. Leyvraz, "Connectivity of growing random networks," *Physical Review Letters*, vol. 85, no. 21, p. 4629, 2000.

[40] P. Holme and B. J. Kim, "Growing scale-free networks with tunable clustering," *Physical Review E*, vol. 65, no. 2, p. 026107, 2002.

[41] K. Klemm and V. M. Eguiluz, "Growing scale-free networks with small-world behavior," *Physical Review E*, vol. 65, no. 5, p. 057102, 2002.

[42] M. Barthélémy, "Crossover from scale-free to spatial networks," *EPL (Europhysics Letters)*, vol. 63, no. 6, p. 915, 2003.

[43] M. Catanzaro, M. Boguná, and R. Pastor-Satorras, "Generation of uncorrelated random scale-free networks," *Physical Review E*, vol. 71, no. 2, p. 027103, 2005.

[44] A. Lancichinetti, S. Fortunato, and F. Radicchi, "Benchmark graphs for testing community detection algorithms," *Physical Review E*, vol. 78, no. 4, p. 046110, 2008.

[45] R. Ibsen-Jensen, K. Chatterjee, and M. A. Nowak, "Computational complexity of ecological and evolutionary spatial dynamics," *Proceedings of the National Academy of Sciences*, vol. 112, no. 51, pp. 15636–15641, 2015.