Solution to urn models of pairwise interaction with application to social, physical, and biological sciences

William Pickering and Chjan Lim
Department of Mathematical Sciences, Rensselaer Polytechnic Institute, 110 8th Street, Troy, New York 12180, USA
(Received 8 February 2017; published 10 July 2017)

We investigate a family of urn models that correspond to one-dimensional random walks with quadratic transition probabilities that have highly diverse applications. Well-known instances of these two-urn models are the Ehrenfest model of molecular diffusion, the voter model of social influence, and the Moran model of population genetics. We also provide a generating function method for diagonalizing the corresponding transition matrix that is valid if and only if the underlying mean density satisfies a linear differential equation and express the eigenvector components as terms of ordinary hypergeometric functions. The nature of the models lead to a natural extension to interaction between agents in a general network topology. We analyze the dynamics on uncorrelated heterogeneous degree sequence networks and relate the convergence times to the moments of the degree sequences for various pairwise interaction mechanisms.

DOI: 10.1103/PhysRevE.96.012311

I. INTRODUCTION

Physical applications of urn problems can be traced to the Ehrenfest model to describe the second law of thermodynamics [1]. However, systems such as the Ehrenfest model do not adequately describe the dynamics of interacting particle systems [2]. To address this, we introduce an extension of the Ehrenfest model that incorporates interactions between two agents. In our models, two balls are randomly drawn from the two urns and are then redistributed stochastically. Naturally, these models have a wide range of physical applications for various interpretations of the urns themselves, such as well-mixed kinetic reactions [3–5] and thermodynamics [1,6]. The Moran model with mutation in population genetics is another member of our rich class of models [7,8]. These two-urn models also have applications to social opinion dynamics, in which the voter model [9–11] is the only case that is a martingale. When the system is generalized to three or more urns, one can analyze Naming Game dynamics on the complete graph [12–16] in a similar fashion. The solutions that we provide here is a generalization of the solution of the Naming Game with many opinions [17]. Further instances of interacting particle systems are the contact process, exclusion processes, and stochastic Ising models [2,9].

In addition to the class of models that describe interacting particle systems, we also provide their exact solutions. As a result, the method solves a large class of random-walk models with quadratic transition probabilities. The method is an extension of the generating function solution of the Ehrenfest model formulated by Mark Kac in 1947 [6]. We utilize a generating function method for solving for all of the eigenvalues and eigenvectors of the Markov transition matrix for each model. With the explicit diagonalization of the transition matrix, we can compute several quantities depending on the application of the model. For instance, in sociophysics [16], the expected time to consensus is one quantity of interest that we can calculate exactly [12,14,18–21]. Based on the parameters of the model, the method of analysis will vary. Figure 1 describes the classification of these models based on solvability and relevant macroscopic properties, which are functions of the model parameters (given in Sec. II).

The method of calculating all eigenvalues and eigenvectors of a transition matrix is complementary to a procedure by Karlin and McGregor [22] to solve for the propagator. In the latter, the procedure requires a tridiagonal transition matrix, whose transition probabilities are strictly positive, which has a consequence that there are no absorbing states in the Markov chain. From such assumptions, authors construct a recursion that yields a polynomial representation of eigenvectors that is orthogonal under a solvable measure. For the bounded models that are proposed here, the recursion terminates with the final condition reducing to the characteristic polynomial of the transition matrix, whose roots become intractable to calculate for large system sizes. To improve this procedure, we focus on a method that includes absorbing states as well as pentadiagonal matrix structure. As such, we primarily consider applications that contain an absorbing state, which also corresponds to a consensus state. Furthermore, the method presented here exactly solves for all eigenvalues without finding roots of a potentially high degree polynomial.

The mechanics of drawing two balls from two urns corresponds to pairwise interaction between agents, which has a natural extension to general networks. In the two-urn model, any combination of two balls can be drawn with equal probability. If we interpret each ball as a node in a network, then the edges correspond to possible pairwise interactions. For the two-urn model, the network is a complete graph, in which every node is connected to every other node. As an extension, we also consider the case in which the network structure is incomplete. That is, after choosing a node randomly, a neighboring node is randomly chosen for pairwise interaction. This is particularly relevant to social interactions and models, such as the voter model and the Naming Game.

The manuscript is structured as follows. In Sec. II, we explicitly define and analyze the two-urn models. We apply a generating function method for exactly diagonalizing a subclass of two-urn models, which we call linear urn models. This solvability distinction is the first split in the diagram given in Fig. 1. In Sec. III, we analyze some of the exact solutions that can be found as a result of the diagonalization. These include the m-step propagator, the moments of consensus time, and local times. In Sec. IV, we analyze the other cases of

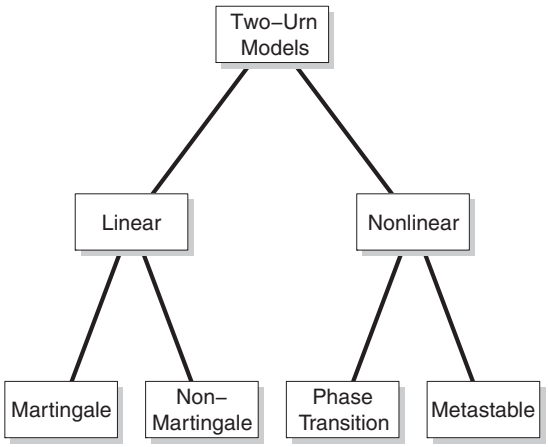

FIG. 1. Classification tree for the two-Urn problems that shows all relevant subclasses. Among the linear cases ($2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 = 0$), the martingales ($\alpha_1 = \beta_1, \alpha_2 = \beta_2 = \gamma_1 = \gamma_2 = 0$) are equivalent to the voter model and the nonmartingales constitute a much larger class of models. Among nonlinear cases, we show that there exists a phase transition over a single parameter, which results in a metastable distribution. The tree shows a general outline of this article prior to the network analysis.

the two-urn models that are nonlinear and therefore cannot be explicitly diagonalized by the generating function method. After assuming that consensus of one urn is stable, we show that there is a phase transition over a single parameter, in which the consensus time abruptly changes from $\ln N$ to exponentially large in N. The latter of these systems tends to a metastable distribution that we solve for exactly. In Sec. V, we consider the case in which equivalent two-urn models are posed on networks and examine the influence of uncorrelated heterogeneous degree network topology on convergence properties. We show that for such a class of networks, the consensus time is exactly equal to the complete graph consensus time multiplied by a single coefficient that describes the network structure.

II. LINEAR TWO-URN MODELS AND SOLUTION

In this section we will define the two-urn models as a stochastically evolving system. We will also define a subclass of the general two-urn framework called linear urn models. For these linear cases, we solve for all eigenvalues and eigenvectors of the transition matrix for the probability distribution of macrostates. With these, we can diagonalize the transition matrix and exactly calculate the probability that the system attains a given macrostate at an arbitrary future time.

In the general model, there are two urns (call them A and B) that have N balls distributed between them. In a discrete time step, two random balls are drawn and redistributed between the urns stochastically. The redistribution probabilities only depend on the urns from which the balls came and the order that they were drawn.

Let $n_A(m)$ denote the number of balls in urn A at discrete time m. The model is characterized by six rate parameters, which we denote by $\{\alpha_1,\alpha_2,\beta_1,\beta_2,\gamma_1,\gamma_2\}$. The parameters are defined in terms of the probability of redistribution between the urns conditioned on which urns the balls came from and their order. These parameters are given to be

$$\alpha_1 = \Pr\{\Delta n_A = 1|AB\} + Pr\{\Delta n_A = 1|BA\}, \quad (1)$$

$$\alpha_2 = \Pr\{\Delta n_A = 1|BB\}, \quad (2)$$

$$\beta_1 = \Pr\{\Delta n_A = -1|AB\} + Pr\{\Delta n_A = -1|BA\}, \quad (3)$$

$$\beta_2 = \Pr\{\Delta n_A = -1|AA\}, \quad (4)$$

$$\gamma_1 = \Pr\{\Delta n_A = 2|BB\}, \quad (5)$$

$$\gamma_2 = \Pr\{\Delta n_A = -2|AA\}. \quad (6)$$

Since these parameters correspond to probabilities, all parameters must be positive, $\alpha_1 + \beta_1 \leqslant 2$, $\alpha_2 + \gamma_1 \leqslant 1$, and $\beta_2 + \gamma_2 \leqslant 1$.

An interpretation of the parameters of the urn model is the influence of agents in social settings. Values of α_1 and β_1 correspond to the impact that an agent has on another with the opposite opinion. Here, two opposing individuals enter a discussion and one of them changes their opinion as a result. This is the case for the voter model, which has parameter configuration $\{1,0,1,0,0,0\}$. The other parameters, α_2, β_2, γ_1, and γ_2 can correspond to mutation and competition between individuals. Also, these parameters can represent push-pull factors to Lee's model of migration [23], and quadratic transition probabilities reflect the assumptions made in Gravity models of migration and trade [24,25]. Existing models with explicit parameter configurations also include the Moran model of genetic drift, with parameters $\{1-\mu_1,\mu_2, 1-\mu_2,\mu_1,0,0\}$, where μ_1 and μ_2 are mutation probabilities [7,8]. For these models, the population size, N, is not always large and thus the discrete stochastic treatment we provide is necessary.

The parameters affect the transition probabilities of the urn model when $n_A = i$, which are given to be

$$p_i^{(1)} = \alpha_1 \frac{i(N-i)}{N(N-1)} + \alpha_2 \frac{(N-i)(N-i-1)}{N(N-1)}, \quad (7)$$

$$p_i^{(2)} = \gamma_1 \frac{(N-i)(N-i-1)}{N(N-1)}, \quad (8)$$

$$q_i^{(1)} = \beta_1 \frac{i(N-i)}{N(N-1)} + \beta_2 \frac{i(i-1)}{N(N-1)}, \quad (9)$$

$$q_i^{(2)} = \gamma_2 \frac{i(i-1)}{N(N-1)}. \quad (10)$$

Here, we define $p_i^{(k)} = \Pr\{\Delta n_A = k|n_A = i\}$ and $q_i^{(k)} = \Pr\{\Delta n_A = -k|n_A = i\}$. Notice that the parameter choice $\{1,1,1,1,0,0\}$ exactly simplifies to the Ehrenfest model. Let $a_i^{(m)} = \Pr\{n_A(m) = i\}$. We introduce the finite difference operator Δ_{ki} acting on a grid function ϕ_i defined as $\Delta_{ki}[\phi_i] = \phi_{i+k} - \phi_i$. We form the single step difference equation that describes the probability distribution in macro-state:

$$a_i^{(m+1)} - a_i^{(m)} = \Delta_{-1i}\left[p_i^{(1)} a_i^{(m)}\right] + \Delta_{-2i}\left[p_i^{(2)} a_i^{(m)}\right] + \Delta_{+1i}\left[q_i^{(1)} a_i^{(m)}\right] + \Delta_{+2i}\left[q_i^{(2)} a_i^{(m)}\right]. \quad (11)$$

This constitutes a pentadiagonal Markov transition matrix for the system. We solve for all eigenvalues and eigenvectors of this model by extending the procedure for diagonalizing the voter model [20]. For eigenvalue λ and eigenvector $\mathbf{v}$ with

components c_i, let $G(x,y) = \sum_i c_i x^i y^{N-i}$ be the generating function for the eigenvectors. We rewrite the spectral problem for the single-step propagator given in Eq. (11) as a partial differential equation for G using the differentiation and shift properties of G [20,26–28]. The partial differential equation for G is

$$\begin{aligned} &N(N-1)(\lambda-1)G \\ &\quad = \gamma_1(x^2-y^2)G_{yy} + \alpha_1 x(x-y)G_{xy} + \alpha_2 y(x-y)G_{yy} \\ &\qquad - \gamma_2(x^2-y^2)G_{xx} - \beta_1 y(x-y)G_{xy} - \beta_2 x(x-y)G_{xx}. \end{aligned} \tag{12}$$

To solve this equation, we make the change of variables $u = x - y$ and $H(u,y) = G(x,y)$. We show below that H has the same structure as G. That is, we define $H(u,y) = \sum_i b_i u^i y^{N-i}$. We make this change of variables because we can solve for b_i and λ exactly when

$$2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 = 0, \tag{13}$$

which we call the linearity constraint. The rationale for relating Eq. (13) with linear systems is given below. If this parameter constraint is not satisfied, then the generating function method does not produce explicit results for the eigenvalues and eigenvectors. Making these substitutions gives the equation for H:

$$\begin{aligned} N(N-1)(\lambda-1)H = {} & [\gamma_1 u^2 + (2\gamma_1 + \alpha_2)uy]H_{yy} + [(\alpha_1 - 2\gamma_1)u^2 + (\alpha_1 - 4\gamma_1 - 2\alpha_2 - \beta_1)uy]H_{uy} \\ & + (\gamma_1 - \alpha_1 - \gamma_2 - \beta_2)u^2 H_{uu}. \end{aligned} \tag{14}$$

Rewriting this as recursion relation for b_i and solving for b_i gives

$$b_i = \frac{\{[(-2\gamma_1 + \alpha_1)(i-1) + (2\gamma_1 + \alpha_2)(N-i)]b_{i-1} + \gamma_1(N-i+2)b_{i-2}\}(N-i+1)}{N(N-1)(\lambda-1) - (\alpha_1 - 4\gamma_1 - 2\alpha_2 - \beta_1)i(N-i) - (\gamma_1 - \alpha_1 - \gamma_2 - \beta_2)i(i-1)}. \tag{15}$$

This allows us to find all eigenvalues exactly. Since $c_i = 0$ for $i < 0$ and $i > N$, we require $b_i = 0$ for $i < 0$ and $i > N$ as well. Since Eq. (15) is an explicit linear difference equation, every $b_i = 0$ unless the equation is singular for some $i = k$. However, this corresponds to the trivial solution to the eigenvalue problem. Thus, the denominator of Eq. (15) must be zero when $i = k$. Solving for λ shows that the eigenvalues are

$$\begin{aligned} \lambda_k = 1 &- \frac{(-\alpha_1 + 4\gamma_1 + 2\alpha_2 + \beta_1)k(N-k)}{N(N-1)} \\ &- \frac{(-\gamma_1 + \alpha_1 + \gamma_2 + \beta_2)k(k-1)}{N(N-1)} \end{aligned} \tag{16}$$

for $k = 0 \ldots N$. This allows b_k to take any value. Values for b_i for any $i > k$ can be found by repeated application of Eq. (15). Expressing $H(u,y)$ in the original coordinates gives

$$G(x,y) = \sum_{i=0}^{N}\sum_{j=i}^{N}(-1)^{j-i}\binom{j}{i}b_j x^i y^{N-i}, \tag{17}$$

which shows that H and G have the same form [20]. Thus the spectral problem is solved for all urn models that satisfy $2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 = 0$. If this equality constraint does not hold, then b_i could not be solved explicitly. This parameter constraint is related to the equation for the mean density of the system. Let $\bar{\rho}$ be the expected value of $\rho = n_A/N$. Using the transition rates, we find that

$$E[\Delta n_A | n_A = i] = 2p_i^{(2)} + p_i^{(1)} - q_i^{(1)} - 2q_i^{(2)}. \tag{18}$$

For $\Delta t = 1/N$, we have that $\Delta n_A = \Delta\rho/\Delta t$. Replacing $i/N \to \bar{\rho}$, we find that for large N, the equation for $\bar{\rho}$ is

$$\begin{aligned} \frac{d\bar{\rho}}{dt} = {} & 2\gamma_1 + \alpha_2 + (\alpha_1 - 4\gamma_1 - 2\alpha_2 - \beta_1)\bar{\rho} \\ & + (2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2)\bar{\rho}^2. \end{aligned} \tag{19}$$

This differential equation is linear if and only if $2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 = 0$, which coincides exactly with our solvability condition for the above generating function method. Therefore, we define a linear urn model as any two-urn model whose parameter configuration yields $2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 = 0$. If this condition is not satisfied, we call the model nonlinear. We conjecture that no other change of variables $(x,y) \to (u,v)$ will solve the nonlinear cases in this fashion, although a proof of this claim is not given. This is the first major categorical distinction shown in Fig. 1.

The treatment of the spectral problem by generating functions is equivalent to a similarity transformation of the transition matrix. Let $\mathbf{T}$ denote the transition matrix given by Eq. (11) and let $\mathbf{w} = \mathbf{P}\mathbf{v}$ for some transformation matrix $\mathbf{P}$. Then, the spectral problem for $\mathbf{w}$ is given by $\mathbf{P}\mathbf{T}\mathbf{P}^{-1}\mathbf{w} = \lambda\mathbf{w}$. The generating function method prescribes the matrix $\mathbf{P}$ so the new matrix $\mathbf{L} = \mathbf{P}\mathbf{T}\mathbf{P}^{-1}$ is lower triangular with a bandwidth of at most two. The transformation, $\mathbf{P}$, is exactly the upper Pascal matrix [29]. This matrix and its inverse is given component-wise by

$$[\mathbf{P}]_{ij} = \binom{j}{i}, \tag{20}$$

$$[\mathbf{P}^{-1}]_{ij} = (-1)^{j-i}\binom{j}{i}. \tag{21}$$

We use the convention that $\binom{j}{i} = 0$ when $i > j$, which implies that $\mathbf{P}$ and $\mathbf{P}^{-1}$ are upper triangular.

The spectral decomposition of the transition matrix can be found by this similarity transformation. We do this by diagonalizing the matrix $\mathbf{L} = \mathbf{W}\Lambda\mathbf{W}^{-1}$. Here, $\Lambda = \mathrm{diag}(\lambda_0, \ldots, \lambda_N)$ and $\mathbf{W}$ are the eigenvectors of $\mathbf{L}$. The components of these eigenvectors are b_i corresponding to

eigenvalue λ_k. Since $b_i = 0$ for $i < j$, $\mathbf{W}$ is lower triangular. Therefore, $\mathbf{W}^{-1}$ can be found explicitly via forward substitution. Diagonalization of $\mathbf{L}$ allows us to explicitly diagonalize the transition matrix as

$$\mathbf{T} = \mathbf{P}^{-1}\mathbf{W}\Lambda\mathbf{W}^{-1}\mathbf{P}. \tag{22}$$

Note that the matrix of eigenvectors is given by $\mathbf{P}^{-1}\mathbf{W}$.

Tridiagonal case: Hypergeometric solution

When the transition matrix for the two-urn model is tridiagonal ($\gamma_1 = \gamma_2 = 0$), we find that the generating function for the eigenvectors can be reduced even further in the above solution. We return to Eq. (15) and consider the following two cases.

1. Case: $\alpha_2 - \alpha_1 \neq 0$

Iterating Eq. (15) when $\gamma_1 = \gamma_2 = 0$ and taking $b_k = 1$ gives

$$b_i = \prod_{j=k+1}^{i} \frac{-\left[j + \frac{\alpha_2 N - \alpha_1}{\alpha_1 - \alpha_2}\right](j - N - 1)}{(j-k)\left[j + k + \frac{(\alpha_2+\beta_2)N - (\alpha_1+\beta_2)}{\alpha_1 - \alpha_2}\right]}, \tag{23}$$

which has explicit solution given by

$$b_i = \frac{(-1)^{i-k}\left(k + \frac{\alpha_2 N - \alpha_2}{\alpha_1 - \alpha_2}\right)_{i-k}(k-N)_{i-k}}{(i-k)!\left[2k + \frac{(\alpha_2+\beta_2)N - (\alpha_2+\beta_2)}{\alpha_1 - \alpha_2}\right]_{i-k}}. \tag{24}$$

Here, the notation $(x)_i$ denotes the Pochhammer symbol [30], which is defined for non-negative integer n by

$$(x)_n = x(x+1)\ldots(x+n-1). \tag{25}$$

For $i < k$, recall that we take $b_i = 0$. Taking these as the coefficients for $H(u,y)$ gives

$$H(u,y) = u^k y^N \sum_{i=0}^{N-k} \frac{(-1)^i\left(k + \frac{\alpha_2 N - \alpha_2}{\alpha_1 - \alpha_2}\right)_i (k-N)_i}{i!\left[2k + \frac{(\alpha_2+\beta_2)N - (\alpha_2+\beta_2)}{\alpha_1 - \alpha_2}\right]_i}\left(\frac{u}{y}\right)^i. \tag{26}$$

Recalling that $u = x - y$ and taking $y = 1$ gives the conventional generating function for the components of the kth eigenvector, given by

$$g_k(x) = G(x,1) = \sum_{i=0}^{N} c_i x^i. \tag{27}$$

Since $G(x,1) = H(x-1,1)$, we have

$$g_k(x) = (x-1)^k {}_2F_1(k + A, k - N; 2k + B; 1 - x), \tag{28}$$

where

$$A = \frac{\alpha_2}{\alpha_1 - \alpha_2}(N-1), \tag{29}$$

$$B = \frac{\alpha_2 + \beta_2}{\alpha_1 - \alpha_2}(N-1). \tag{30}$$

2. Case: $\alpha_1 = \alpha_2$

This case corresponds to the generalization of the Ehrenfest urn model by Krafft and Schaeffer [31,32]. This is because we assume $\gamma_1 = \gamma_2 = 0$, $\alpha_1 = \alpha_2 = \alpha$, and therefore $\beta_1 = \beta_2 = \beta$ by the linearity constraint. Taking Eq. (15) with these parameters gives

$$b_i = \prod_{j=k+1}^{i} \frac{-\alpha(j-N-1)}{(\alpha+\beta)(j-k)}, \tag{31}$$

$$= \left(-\frac{\alpha}{\alpha+\beta}\right)^{i-k} \frac{(k-N)_{i-k}}{(i-k)!}, \tag{32}$$

$$= \left(\frac{\alpha}{\alpha+\beta}\right)^{i-k} \binom{N-k}{i-k}. \tag{33}$$

Therefore, H is given by

$$H(u,y) = u^k y^N \sum_{i=0}^{N-k} \binom{N-k}{i}\left[\frac{\alpha u}{(\alpha+\beta)y}\right]^i, \tag{34}$$

$$= u^k y^N \left[\frac{\alpha u}{(\alpha+\beta)y} + 1\right]^{N-k}. \tag{35}$$

Taking $g_k(x) = G(x,1) = H(x-1,1)$ as the conventional generating function for the eigenvector components gives

$$g_k(x) = (x-1)^k \left(\frac{\alpha x + \beta}{\alpha + \beta}\right)^{N-k}, \tag{36}$$

which is a generalization of the solution to the Ehrenfest model given by Kac [6].

III. APPLICATIONS OF THE SPECTRAL SOLUTION

The immediate consequence of the explicit diagonalization of the transition matrix is the solution of the m step propagator. That is, we can find all future probability distributions of macrostates explicitly. Let d_k be the initial distribution expressed in the eigenbasis. That is, d_k are the components of $\mathbf{d} = \mathbf{W}^{-1}\mathbf{P}\mathbf{a}^{(0)}$. With d_k known, all future probability distributions are given by

$$a_i^{(m)} = \sum_{k=0}^{N} d_k \lambda_k^m [\mathbf{v}_k]_i. \tag{37}$$

Using this, other exact solutions can be found, such as the expected time to consensus and expected local times. These solutions are summarized in Table I and the details are given in the following subsections.

A. Consensus time

The consensus time is the amount of time spent before the system reaches an absorbing state in which all of the balls are in a single urn. In this section, we only consider linear models in which consensus of B is absorbing. This is because if both consensus of A and B were absorbing, then $\alpha_2 = \beta_2 = \gamma_1 = \gamma_2 =$

TABLE I. Summary of exact solutions.

Quantity	Discrete Solution
Macro-state probability	$a_i^{(m)} = \sum_{k=0}^{N} d_k [\mathbf{v}_k]_i \lambda_k^m$
Consensus time	$E[\tau^p] \sim \sum_{k=1}^{N} \frac{d_k p!}{[N(1-\lambda_k)]^{p+1}}$ $\times\left\{\beta_1[\mathbf{v}_k]_1 + \frac{2\gamma_2}{N-1}[\mathbf{v}_k]_2\right\}$
Local time	$E[\mathbf{M}] \sim \frac{1}{N}\sum_{k=0\lambda_k\neq 1}^{N} \frac{d_k}{1-\lambda_k}\mathbf{v}_k$

0. By linearity, we have $\alpha_1 = \beta_1$, which reduces to the voter model on the complete graph. These are the only cases of the two-urn models that are martingale, and has already been addressed in previous work [20]. Therefore, we only consider the nonmartingale cases in which consensus of one urn is absorbing. For $n_A = 0$ to be an absorbing state, we require $\Pr\{\Delta n_A = 0 | n_A = 0\} = 1$. For this to be true, we need $\alpha_2 = \gamma_1 = 0$.

We can find the moments of the consensus time by using the solution to the spectral problem. Let s_m be the probability that the system reaches consensus at discrete time m. Let $\tau = m/N$ be the consensus time. Then, we have

$$E[\tau^p] = \sum_{m=1}^{\infty} s_m (m/N)^p. \tag{38}$$

The probability of consensus, s_m, can be expressed as the probability that the system is one step from consensus multiplies by the probability that the system moves into the consensus state. This is expressed as $s_m = q_1^{(1)} a_1^{(m-1)} + q_2^{(2)} a_2^{(m-1)}$. Since $a_i^{(m)}$ is given by the diagonalization of the transition matrix, we write this as

$$s_m = \sum_{k=1}^{N} d_k \lambda_k^{m-1} \left\{ \frac{\beta_1}{N} [\mathbf{v}_k]_1 + \frac{2\gamma_2}{N-1} [\mathbf{v}_k]_2 \right\}. \tag{39}$$

Making this substitution into Eq. (38) and evaluating the infinite geometric series gives

$$E[\tau^p] \sim \sum_{k=1}^{N} \frac{d_k\, p!}{[N(1-\lambda_k)]^{p+1}} \left\{ \frac{\beta_1}{N} [\mathbf{v}_k]_1 + \frac{2\gamma_2}{N-1} [\mathbf{v}_k]_2 \right\}. \tag{40}$$

This solution for the moments of consensus time is exact; however, it does not clearly describe how it depends on N asymptotically. We can approximate the consensus time without resorting to the full diagonalization of the transition matrix. We do this by estimating when the survival probability of the system becomes small $(1/N)$. With high probability, the drift moves the system near consensus in $O(1)$ time. Then, the probability that the system has not reached consensus by discrete time m can be estimated by λ_1^m. Since we define the consensus time to be $\tau = m/N$, we wish to solve

$$\lambda_1^{\tau N} = \frac{1}{N}. \tag{41}$$

Using the eigenvalues given by Eq. (16), we have that $\lambda_1 = 1 - (\beta_1 - \alpha_1)/N$. Therefore, solving for τ and using $\ln \lambda_1 \sim -(1 - \lambda_1)$, we find that

$$E[\tau] \sim \frac{\ln N}{\beta_1 - \alpha_1} + O(1). \tag{42}$$

Figure 2 compares simulation data against this estimate and shows that there is good agreement between them.

B. Local time

The local time is the amount of scaled time (m/N) that the system spends at each macrostate prior to consensus. Again, we assume that consensus of B is the only absorbing consensus state. The local time is given as a vector $\mathbf{M}$, whose components, M_i, are the local time for the corresponding macrostate. Let

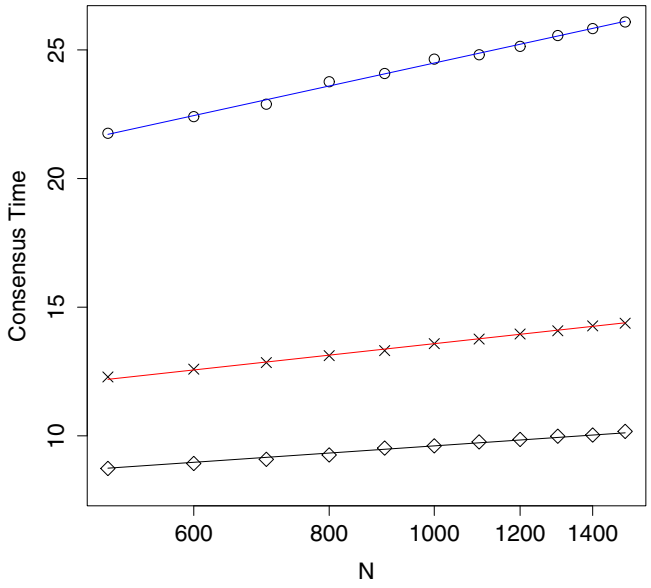

FIG. 2. Consensus times for three linear models averaged over 3 000 runs is plotted on a logarithmic scale in N. These models have parameter configurations {1/4,0,1,1/4,0,1/4} (○), {1/2,0,1,1/6,0,1/6} (×), and {3/4, 1,1/2,0,1/12} (◇). Solid lines are the estimates given by Eq. (42) with a fitted additive constant. The constant is dominated by the logarithm for large N.

$M_i(m)$ be the local time for macrostate $n_A = i$ by time m. To find the expected local time, we consider $E[\Delta M_i(m)]$, where $\Delta M_i(m) = M_i(m+1) - M_i(m)$. The change in the local time from step m to $m+1$ is $1/N$ if the system is in macrostate $n_A = i$ and 0 otherwise. So, the expected change in the local time is equal to $a_i^{(m+1)}$. This implies that the expected time spent at a macrostate prior to consensus is expressed as

$$E[M_i(\infty)] - E[M_i(0)] = \sum_{m=0}^{\infty} E[\Delta M_i(m)] = \frac{1}{N} \sum_{m=0}^{\infty} a_i^{(m+1)}. \tag{43}$$

$E[M_i(0)]$ is exactly the initial probability distribution of the model: $a_i^{(0)}$. Using the diagonalization of linear urn models given by Eq. (37), we can find $E[M_i(\infty)]$ exactly. Making the substitution, evaluating the geometric series gives, and organizing $M_i(\infty)$ into the vector $\mathbf{M}$ gives

$$E[\mathbf{M}] = \frac{1}{N} \sum_{k=1}^{N} \frac{d_k \mathbf{v_k}}{1 - \lambda_k}. \tag{44}$$

The $k = 0$ term is not included because $d_0 = 0$. This is because $k = 0$ corresponds to the consensus state, and we stop measuring the local time once consensus has been attained.

IV. NONLINEAR MODELS

Here we study models for which $2\gamma_1 - 2\gamma_2 - \alpha_1 + \alpha_2 + \beta_1 - \beta_2 \neq 0$. This means that the eigenvalues and eigenvectors of the transition matrix cannot be explicitly found by the generating function method proposed in Sec. II. We show below that nonlinear models also have a much wider range of qualitative behaviors that would be difficult to extract from a

similar generating function method. In particular, we consider nonlinear models such that consensus of B is an absorbing state. First, we will show that there exists a phase transition across $\alpha_1 = \beta_1$.

A. Phase transition

Here we consider similar stable consensus type models as in the linear case. That is, when all of the balls are in urn B, then all of them will remain in urn B with probability 1. Consensus of B is an absorbing state in the Markov-chain model. For this to be true, we must set $\gamma_1 = \alpha_2 = 0$. We take Eq. (19) and make these substitutions to find

$$\frac{d\bar{\rho}}{dt} = (-2\gamma_2 - \alpha_1 + \beta_1 - \beta_2)\bar{\rho} \times \left(\frac{\alpha_1 - \beta_1}{-2\gamma_2 - \alpha_1 + \beta_1 - \beta_2} + \bar{\rho}\right). \tag{45}$$

Clearly, $\bar{\rho}_1 = 0$ is stationary. This is anticipated because consensus of B is an absorbing state. However, this root is not always stable and that there exists a phase transition when $\alpha_1 = \beta_1$. This can be seen by considering Eq. (45) for $\bar{\rho} \ll 1$.

The second root, $\bar{\rho}_2$, can be expressed by a single parameter, ω. We let

$$\omega = \frac{\alpha_1 - \beta_1}{2\gamma_2 + \beta_2} \tag{46}$$

and observe that

$$\bar{\rho}_2 = \frac{\omega}{1+\omega}. \tag{47}$$

The location and stability of the roots are characterized completely by ω because for $\omega > 0$, the root at the origin is unstable and $\bar{\rho}_2$ becomes stable. The root ρ_2 also exists within the physical domain ($0 \leqslant \bar{\rho} \leqslant 1$), which indicates that the system is attracted to a metastable state. Stochastically, the system will also randomly fluctuate within this interval and always has a nonzero probability of achieving consensus. So, even though the system is attracted to the metastable distribution with high probability, it will eventually achieve the consensus state in finite time. The consensus time for the metastable case, however, is exponential with N. The special case when $\gamma_2 = \beta_2 = 0$ should be interpreted as $\omega = \pm\infty$ and corresponds to a biased voter model [19]. In this case, the consensus time may not be exponential because the stable equilibrium is a consensus state.

For $\omega < 0$, the origin is stable, and ρ_2 is unstable. In this case, the solution to Eq. (45) is the logistic function. For $\bar{\rho}$ small, the equation resembles a linear model, so the consensus time is qualitatively the same as the linear cases. That is, the consensus time is $E[\tau] \sim \ln N/(\beta_1 - \alpha_1)$. So, we will only focus on the remaining cases: the phase transition itself ($\omega = 0$) and the metastable state ($\omega > 0$).

When $\omega = 0$, the two roots are coincident, which indicates a transcritical bifurcation [33]. Figure 3 shows the bifurcation diagram of these nonlinear two-urn models.

Although the consensus time is proportional to $\ln N$ when $\omega < 0$, we will show that when $\omega = 0$, this value increases to $\sqrt{N}$ for $\gamma_2 + \beta_2 \neq 0$.

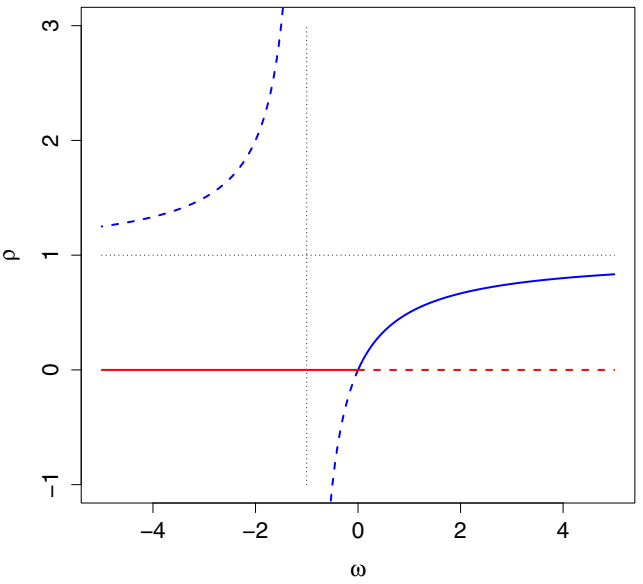

FIG. 3. Bifurcation diagram showing the stability of the stationary points of Eq. (45) for nonlinear consensus models. The solid lines indicate that the point is stable, while dashed lines indicate that it is unstable. The red lines correspond to the root at $\bar{\rho}_1 = 0$, and the blue curve corresponds to the second root, $\bar{\rho}_2$. For $\omega > 0$, the system attracts to a metastable state and for $\omega < 0$, the systems resemble linear models.

Consensus formation with mutation

Socially, we consider a model that may correspond to the decay of a popular trend, A, that dominates the system. Those who do are not involved with the trending fad are considered to have state B. Those who do not accept A can be convinced to adopt it with probability α_1 when exposed to it. Similarly, people with A can be dissuaded from it with probability β_1 when exposed to agents without A. Also, an individual may be turned away from the fad if they are exposed to it for long enough. That is, if an individual with A is exposed to many individuals with A, then the trend may seem less unique, and the agent no longer adopts it, which is a sociologically observed tendency [34]. If A is perceived as a fad, then people will be less likely to adopt it [35]. We model this by assuming that if A speaks with another A, then the fad is rejected with probability β_2 and the listener becomes B. It may also be reasonable to assume the both individuals would reject the fad simultaneously, which would occur with probability γ_2, but, for simplicity, we set $\gamma_2 = 0$.

We wish to find the amount of time until nobody accepts A, and everyone is in state B. To simplify the analysis, we will consider the case in which $\alpha_1 = \beta_1 = \beta_2 = 1$, although any configuration of parameters could be considered, provided $\omega = 0$. Let $T(\rho)$ be the time to consensus. Using first step analysis, we find that T satisfies

$$T(\rho) = p(\rho)T(\rho + \Delta\rho) + r(\rho)T(\rho) + q(\rho)T(\rho - \Delta\rho) + \frac{1}{N}, \tag{48}$$

where $\Delta\rho = 1/N$ and

$$p(\rho) = \rho(1-\rho), \tag{49}$$

$$q(\rho) = \rho(1-\rho) + \rho^2, \tag{50}$$

$$r(\rho) = 1 - p(\rho) - q(\rho). \tag{51}$$

We also have the boundary condition $T(0) = 0$. Expanding by Taylor's theorem to second order gives

$$-1 \sim v(\rho)\frac{dT}{d\rho} + \frac{1}{2N}D(\rho)\frac{d^2T}{d\rho^2}, \tag{52}$$

where

$$v(\rho) = -\rho^2, \tag{53}$$

$$D(\rho) = 2\rho(1-\rho) + \rho^2. \tag{54}$$

The system is dominated by the drift term when the system is not near consensus. Therefore, ρ approaches consensus in $O(1)$ time. After this time has passed, we assume that these two terms in Eq. (52) balance. Let $\rho = \delta\xi$ for some $\delta \ll 1$ that balances the drift and diffusion terms. The derivatives are then transformed by

$$\frac{d}{d\rho} = \frac{1}{\delta}\frac{d}{d\xi}. \tag{55}$$

So, after dropping higher-order terms, Eq. (52) is given by

$$-1 \sim -\delta\xi^2\frac{dT}{d\xi} + \frac{\xi}{N\delta}\frac{d^2T}{d\xi^2}. \tag{56}$$

For these terms to balance, we choose $\delta = 1/\sqrt{N}$. Now, we wish to solve

$$\xi\frac{d^2T}{d\xi^2} - \xi^2\frac{dT}{d\xi} = -\sqrt{N}. \tag{57}$$

We are given one boundary condition, $T(0) = 0$, yet we wish to solve a second-order equation. Since the time for the system to approach consensus is $O(1)$ and that the consensus time is monotonic, the derivative of consensus time with respect to ρ is at most $O(1)$. So, the derivative of T with respect to ξ tends to zero. This is stated mathematically by

$$\lim_{\xi\to\infty}\frac{dT}{d\xi} = 0, \tag{58}$$

which supplies the second boundary condition. Now, the solution to Eq. (57) is

$$T(\xi) = \sqrt{N}\int_0^{\xi}\int_u^{\infty}\frac{1}{s}e^{-\frac{1}{2}(s^2-u^2)}dsdu. \tag{59}$$

We are particularly interested in the case when A initially dominates the system. Furthermore, since the integral converges exponentially, we can take $\xi \to \infty$ without significantly changing the value. So, to find $T(\infty)$, we take

$$s = \sqrt{2}r\sec\theta, \tag{60}$$

$$u = \sqrt{2}r\tan\theta. \tag{61}$$

Making this change of variables, the integral can be simplified to

$$T \sim \sqrt{\frac{\pi^3 N}{8}}. \tag{62}$$

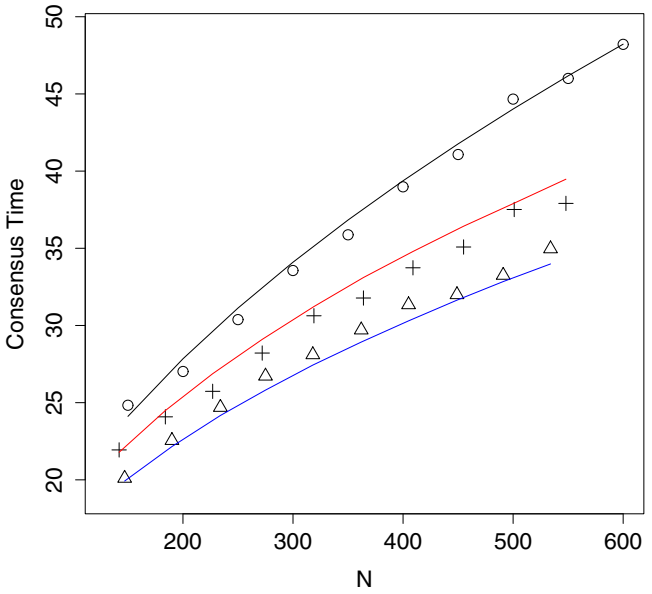

FIG. 4. Solutions for the consensus time given by Eq. (62) compared with simulation data. The lines are the theoretical prediction and the points are simulation data averaged over 1000 runs. Three networks are considered: the complete graph (black, ○) and two scale free networks with $\nu = 1$ (blue, △) and $\nu = 2$ (red, +).

This result is used extensively when analyzing heterogeneous networks in Sec. V. Figure 4 shows the agreement between Eq. (62) and simulation data for the complete graph as well as scale-free networks.

B. Metastable consensus time

The other interesting case is when $\omega > 0$, in which there is a stable fixed point of the drift equation in the feasible region $0 \leqslant \rho \leqslant 1$. Even though the deterministic part of the system is attracted to ρ_2 with high probability, it is not the absorbing state of the system. Consensus of B is stable with probability 1, yet ρ_2 is stable only with probability close to 1. Therefore, ρ_2 is the mean of a metastable distribution that exists apart from consensus. In terms of the Markov chain model, consensus corresponds to the eigenvalue 1 of the transition matrix and the metastable distribution is an eigenvector with eigenvalue that is transcendentally close to 1.

We have two goals of this section. The first is to find the metastable distribution for a particular metastable urn process and the second is to find the corresponding consensus time. As in the above case, we restrict our study to a particular parameter configuration to serve as a canonical example of the solution. Once the balls have been selected from the urns, the model that we consider is given by the following rules:

(1) If the balls came from different urns, then place both in A.

(2) If the balls came from the same urn, then place both in B.

It is evident from these rules that consensus of B is an absorbing state. If all balls are in urn B, then the balls are always drawn from the same urn and replaced in B with probability 1. This model has the parameter configuration

$\alpha_1 = 2$, $\gamma_2 = 1$, and all others equal to 0. By Eq. (46), we have that $\omega = 1$ and, therefore, $\rho_2 = 1/2$. Since $\omega > 0$, the bifurcation diagram in Fig. 3 indicates that $\rho_2 = 1/2$ is stable and constitutes a metastable state.

Let λ be the eigenvalue that corresponds to the metastable distribution, ϕ_i. Since λ is exponentially close to 1, $1 - \lambda$ is a transcendentally small term. Since ϕ_i constitutes components of a eigenvector, we can choose how scale it. Conditioned on the information that the system is not in consensus, then $\phi_i = \Pr\{n_A = i\}$. So, for ϕ_i to constitute a probability distribution, we need the sum of the components to be 1. So, we let c_i be the components of the unscaled eigenvector corresponding to λ and scale it so

$$\phi_i = \frac{c_i}{\sum_{i=1}^{N} c_i}. \tag{63}$$

With this, we obtain a recursion relation for c_i from the eigenvalue problem for c_i and λ:

$$0 = p_{i-1}^{(1)} c_{i-1} - \left(p_i^{(1)} + q_i^{(2)}\right) c_i + q_{i+2}^{(2)} c_{i+2}. \tag{64}$$

We also require that $c_i = 0$ when $i < 0$ and $i > N$. Also, we set $c_N = 1$ to begin the recursion. We know that $c_N \neq 0$ because if $c_N = 0$, it would imply that the system can never achieve this state, which is not true. So, we take Eq. (64) and solve for c_{i-1} to find

$$c_{i-1} = \frac{\left(p_i^{(1)} + q_i^{(2)}\right) c_i - q_{i+2}^{(2)} c_{i+2}}{p_{i-1}^{(1)}}. \tag{65}$$

Given $c_N = 1$ and $c_i = 0$ for $i > N$, we can use Eq. (65) to find c_i for $1 \leqslant i \leqslant N - 1$. This yields a distribution that is asymptotically a Gaussian and is shown in Fig. 5.

Now that we found the metastable distribution, our second goal was to find the consensus time. The strategy is to use Eq. (40), which gives the consensus time in terms of the eigenvalues and eigenvectors. When using Eq. (40), we take $d_1 = 1$ and all other $d_k = 0$ for $[\mathbf{v_1}]_i = \phi_i$ and λ_1 being the eigenvalue for ϕ_i. That is, Eq. (40) reduces to

$$E[\tau] = \frac{2\phi_2}{N^2(N-1)(1-\lambda)^2}. \tag{66}$$

Since we found the dominant eigenvector in the system, all we need is the corresponding eigenvalue. We find it by considering the column-sum of the transition matrix $\mathbf{T}$. Since all of the columns sum to 1, we multiply both sides of $\mathbf{Tv} = \lambda\mathbf{v}$ by $\mathbf{1}^T$ on the left to obtain $\mathbf{1}^T\mathbf{v} = \lambda\mathbf{1}^T\mathbf{v}$. Unless $\lambda = 1$, we must have that the sum of the components of $\mathbf{v}$ is zero. This is important because $\sum_{i=1}^{N} \phi_i = 1$ in order to be a probability distribution conditioned that the system has not reached consensus. This implies that $\phi_0 = -1$ in order for the sum of the components from $i = 0 \ldots N$ is zero. If we take $i = 0$ in the eigenvalue problem, then we find that

$$\lambda\phi_0 = \phi_0 + q_2^{(2)}\phi_2. \tag{67}$$

For $\phi_0 = -1$, we have that $1 - \lambda = q_2^{(2)}\phi_2$. Substituting this into Eq. (66) gives

$$E[\tau] = \frac{N-1}{2\phi_2}, \tag{68}$$

where ϕ_2 is easily found by (65) and (63). Figure 6 compares simulation data against the results in Eq. (68).

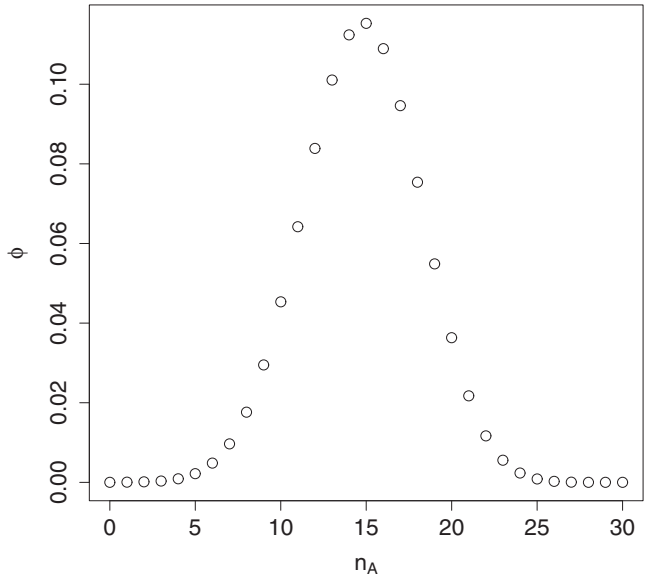

FIG. 5. Solution for the metastable distribution generated by Eq. (65) for $N = 30$. The distribution is asymptotically a Gaussian function because the diffusivity is smooth. Near $\rho_2 = 1/2$, which is the mean, the diffusivity can be approximated by a constant. This yields a Gaussian to leading order.

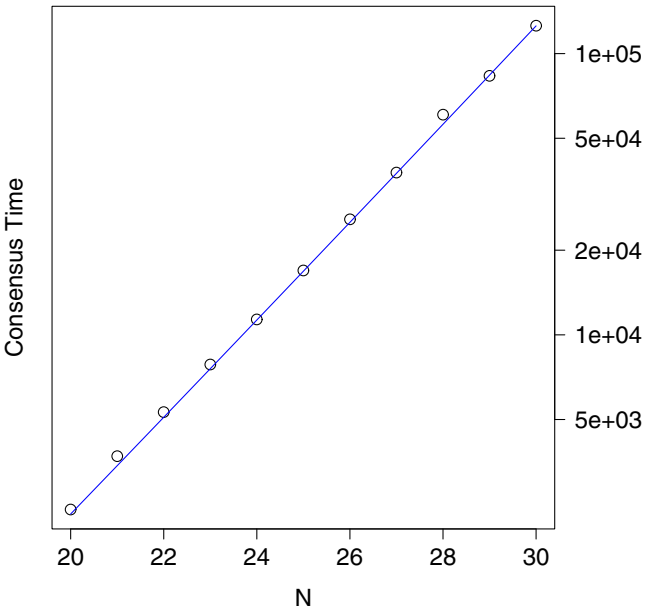

FIG. 6. Simulation data (○) of the consensus time averaged over 500 runs is plotted against the exact solution shown in blue, given by Eq. (68). The horizontal axis is on a linear scale in N and the vertical axis is on a logarithmic scale. The apparent linear relationship indicates the exponential nature of the consensus time in this case.

V. HETEROGENEOUS DEGREE NETWORK MODELS

In the above analysis, we considered social models exclusively on the complete graph. It has been observed that the ordering dynamics of the Naming Game—a similar

TABLE II. Parameters of the network model.

Input				
AA	AB	BA	BB	Output
–	α_{12}	α_{11}	γ_1	AA
β_{22}	–	r_0	α_{21}	AB
β_{21}	r_1	–	α_{22}	BA
γ_2	β_{11}	β_{12}	–	BB

social model—do not significantly change when comparing real-world networks against complete graphs [36,37]. This yields credibility to complete graph analysis for some sparse real-world networks. However, the use of a fully connected graph may not be an accurate approximation of the dynamics on sparse networks, which are also more socially realistic. For the voter model, which is a case of the above general social framework, the system orders differently based on network structure [18–20,38].

The structure of the two-urn models yields a natural extension to incomplete graphs. Each node has one of two states: A or B. For the general network model, we choose a node randomly and then choose a random neighbor. The nodes then update their spins with probabilities that depend on the order that the nodes were selected and their spins. We now have 12 parameters that characterize each model, each of which are transition probabilities. Table II shows the probability of accepting a particular update as a function of the spins and their order. So, if the first node has state B and the second node has state A, then the probability that they both become A is α_{11}. For parameters that have two subscripts (e.g., α_{11}), the second digit refers to which node is updated. The first digit is for consistency with the complete graph (two-urn) notation.

These models have a very wide range of applications. For instance, the voter model has $\alpha_{11} = \beta_{11} = 1$ and all others zero, while the invasion process has $\alpha_{12} = \beta_{12} = 1$, and all others equal to zero [19]. Generalized versions of the Moran model have parameters $\beta_{21} = \mu_1$, $\alpha_{21} = \mu_2$, $\beta_{11} = 1 - \mu_2$, and $\alpha_{11} = 1 - \mu_1$ [39]. We can also cast movement of individuals, or particles, between sites by interpreting state A as occupied and B as unoccupied. Of particular interest in the literature is the coalescing random walk, in which particles move from site to site and cluster together when moving into an occupied site [21,38,40]. This is accomplished in our framework by $r_1 = 1, \beta_{21} = 1$ and all other parameters set to zero. The annihilating random walk is similar to the coalescing random walk but is characterized by the fact that the two particles are removed if one moves into an occupied site [40,41]. This is given by $r_1 = 1, \gamma_2 = 1$ and all other parameters set to zero. Furthermore, we can consider directed networks which determines the flow of information throughout the network. These represent a mere fraction of the possible applications of this general network model.

Clearly, not all parameter configurations on all networks can be solved as swiftly and concisely as the above two-urn models. Even a generalized treatment of each case is beyond the scope of this article. However, we explore some of the previous examples imposed on a popular class of networks that we call uncorrelated degree heterogeneous graphs. These are random networks that are generated with a given, fixed degree sequence. The network is then chosen uniformly from the set of all graphs with that degree sequence [19,42–45]. In the following subsections, we solve two models of consensus formation, which are both closely related to their complete graph counterpart.

To find ordering dynamics, we begin by following the procedure prescribed in Refs. [19,46] to coarse-grain the system and look for solutions in the mean field of the network. This is done by considering the number of nodes with state A that have degree k and averaging over all adjacency matrices with a given degree sequence that are also uncorrelated and heterogeneous. For the voter model, Refs. [19,46] indicates that the system reduces to a one-dimensional diffusion process. However, in the general case of the two-urn models on heterogeneous networks that we consider, the process does not reduce to a simple diffusion process. Despite this, we present a perturbation technique to solve for expected times to consensus, which can be applied whenever consensus is an absorbing state in the Markov chain. We proceed by identifying the coarse-grained system as an urn model, similarly to the two-urn models of except there are more than two urns as in Sec. II. In the new urn model framework, each degree value in the network corresponds to a pair of urns, call them A_k and B_k. This framework allows us to relate the network model to nonlinear two-urn models, for which we directly apply the solutions given in Sec. IV.

A. Consensus model: Listener first

Here we explore a model of consensus formation on uncorrelated heterogeneous networks. We interpret the situation in the following way. During a single update, the first node that is chosen is called the listener. This individual chooses one of its neighbors to be the second node, called the speaker. Then, we assume that only the listener is allowed to update its state. We also assume that consensus of the B state is an absorbing macrostate. That is, if every node is in state B, the every node will remain in state B with probability 1.

Now we express these ideas as a parameter configuration in the above framework. Since only the first node is allowed to change its state in an update, we set $\alpha_{12} = \gamma_1 = \beta_{22} = r_0 = r_1 = \alpha_{22} = \gamma_2 = \beta_{12} = 0$. By assumption, consensus of B is an absorbing state, so we additionally take $\alpha_{21} = 0$. This leaves us with parameters $\alpha_{11}, \beta_{21}, \beta_{11}$. We then drop the second digit in the subscripts since it is implied that only the first node is allowed to change. This corresponds to a three-parameter system: α_1, β_2, β_1.

We wish to express the transition probabilities of the system in terms of known quantities. First, we must establish the notation used in the following discussion. Let $\mathbf{A}$ be the adjacency matrix of the network and let N_k be the number of nodes with degree k. Also let k_i be the degree of node i and let $\mathbf{k}$ be a vector with components k_i. Let $\eta_i = 1$ if node i has state A and $\eta_i = 0$ if node i has state B and is the microstate at time t. The vector $\boldsymbol{\eta}$ takes components η_i. Similarly, let n_k be the number of nodes of degree k of state A and let $\mathbf{n}$ take components n_k. Also, let $\rho_k = n_k/N_k$ and $\rho = \sum_k n_k/N$. These are all related functions of the microstate. Let D_k be the

set of all nodes with degree k. Let μ_p be the pth moment of the degree distribution of the network, which is given by

$$\mu_p = \sum_k \frac{N_k}{N} k^p. \tag{69}$$

Also, we define z_p to be the expected value of $\boldsymbol{\eta}$ with probability measure proportional to k_i^p. We could also define z_p as the normalized instantaneous average value of the p-degree of A nodes. This is written explicitly by

$$z_p = \frac{\mathbf{k}^p \cdot \boldsymbol{\eta}}{N\mu_p}. \tag{70}$$

Here, $\mathbf{k}^p$ is the component-wise power of $\mathbf{k}$. That is, the ith component of $\mathbf{k}^p$ is k_i^p. Finally, we can express the transition probabilities of the network model:

$$\Pr\{\Delta n_k = 1\} = \sum_{i\in D_k}\sum_j \frac{\alpha_1 A_{ij}}{N k_i}(1-\eta_i)\eta_j, \tag{71}$$

$$\Pr\{\Delta n_k = -1\} = \sum_{i\in D_k}\sum_j \frac{A_{ij}}{N k_i}[\beta_1\eta_i(1-\eta_j) + \beta_2\eta_i\eta_j]. \tag{72}$$

In order to find the consensus time, we replace A_{ij} with mean adjacency matrix of the uncorrelated degree heterogeneous networks [19]. Note that we are interested in analyzing the dynamics over an ensemble of random networks in mean field, and the connection of the network is independent of the urn-process. We also show numerically in Fig. 7 that averaging trajectories over an ensemble of networks agrees with this mean-field assumption. Since the edges of the network are uncorrelated, the probability that nodes i and j have an edge is proportional to their respective degrees. The expected value of the components of the adjacency matrix, therefore, is

$$E[A_{ij}] = \frac{k_i k_j}{N\mu_1}. \tag{73}$$

On substitution of $A_{ij} \to E[A_{ij}]$ and simplification, we express the coarse-grained transition probabilities below in terms of the products ρ_k and z_1,

$$p(\rho_k, z_1) = \alpha_1 \frac{N_k}{N}(1-\rho_k)z_1, \tag{74}$$

$$q(\rho_k, z_1) = \beta_1 \frac{N_k}{N}\rho_k(1-z_1) + \beta_2 \frac{N_k}{N}\rho_k z_1. \tag{75}$$

What is significant here is that a single variable $z_1(t)$—the normalized average degree of the A nodes at time t—plays the role of mean-field macrostate (over all degrees k) and allows the coarse-grained transition probabilities to be expressed component-wise in k as binary (quadratic) interactions between the kth macrostate and z_1. This also allows us cast the system as an urn model in which each degree k corresponds to a pair of urns (A_k and B_k). Here n_k would correspond to the number of balls in urn A_k. The movement of balls among these urns is coupled only by z_1.

We use the transition probabilities to generate an ordinary differential equation (ODE) for the time evolution of the expected value of ρ_k. Taking $\dot{\bar{\rho}}_k \sim E[\Delta\rho_k/(1/N)]$, we acquire the following ODE system:

$$\dot{\bar{\rho}}_k = [\alpha_1(1-\bar{\rho}_k) - \beta_2\bar{\rho}_k]z_1 - \beta_1\bar{\rho}_k(1-z_1). \tag{76}$$

Note that this system is coupled together only by the variable z_1. We will first show that $\rho_k \to z_1$ as the system evolves. This is found by expressing the equation for $\dot{\bar{\rho}}_k - \dot{z}_1$. To do so, we need $\dot{z}_1$. Multiplying Eq. (76) by $\frac{N_k}{N\mu_1}$ and summing over k gives

$$\dot{z}_1 = [\alpha_1(1-z_1) - \beta_2 z_1]z_1 - \beta_1 z_1(1-z_1). \tag{77}$$

Note that Eq. (77) can be solved, with solution

$$z_1(t) = \frac{\omega(1+\omega)^{-1}}{1 + \left(\frac{\omega}{(1+\omega)z_1(0)} - 1\right)e^{-(\alpha_1+\beta_1)t}}, \tag{78}$$

where ω is given in Eq, (46). Note that when $\omega > 0$, $z_1 \to 0$ as $t \to \infty$. When $\omega < 0$, we have $z_1 \to \omega(1+\omega)^{-1}$ instead. This solution holds only if $\omega \neq 0$. The point $\omega = 0$ is a bifurcation point, in which the solution is

$$z_1(t) = [\beta_2 t + z_1(0)^{-1}]^{-1}. \tag{79}$$

Once we show that each $\rho_k \to z_1$, we will have explicit solutions of each ρ_k. Now $\dot{\bar{\rho}}_k - \dot{z}_1$ can be expressed as

$$\dot{\bar{\rho}}_k - \dot{z}_1 = -(\bar{\rho}_k - z_1)[(\alpha_1+\beta_2)z_1 + \beta_1(1-z_1)]. \tag{80}$$

Note that the expression in the bracket of Eq (80) is always positive and bounded away from 0. Therefore, Eq. (80) shows an exponential convergence of $\bar{\rho}_k$ to z_1 for each k. This principle is depicted in Fig. 7.

Now that we understand how the system orders, we now calculate the consensus time, $T(\boldsymbol{\rho})$, where $\boldsymbol{\rho}$ takes components ρ_k. The procedure is to relate the problem to a complete graph and then use previously established techniques. For simplicity, we take $\alpha_1 = \beta_1 = \beta_2 = 1$. Note that Eq. (77) converges to

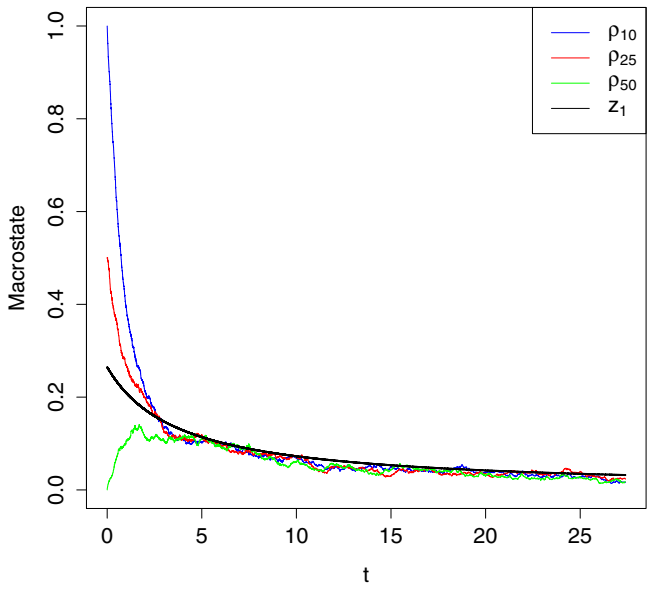

FIG. 7. The listener-only consensus model is simulated for a network with $N = 3000$ with $N_{10} = N_{25} = N_{50} = 1000$. The model was simulated with $\alpha_1 = \beta_1 = \beta_2 = 1$. The initial condition is $\rho_{10} = 1$, $\rho_{25} = 0.5$, and $\rho_{50} = 0$. Also plotted is the exact solution for z_1 given by Eq. (79). This shows that each ρ_k converges to z_1 as the system evolves.

a stable center manifold near $z_1 = 0$. Because of this, the noise terms will be significant near $z_1 \sim 0$. Therefore, we incorporate these terms in our analysis of the consensus time. We first establish the backwards equation for T:

$$-\frac{1}{N} = \sum_k \{q(\rho_k,z_1)T(\rho_k - 1/N_k) + [-q(\rho_k,z_1) - p(\rho_k,z_1)]T(\rho_k) + p(\rho_k,z_1)T(\rho_k + 1/N_k)\}. \quad (81)$$

We expand T to two terms and express the backwards equation as

$$-1 = \sum_k \left(v_k \frac{\partial T}{\partial \rho_k} + \frac{D_k}{2N_k}\frac{\partial^2 T}{\partial \rho_k^2} \right), \quad (82)$$

where

$$v_k = (1-\rho_k)z_1 - \rho_k(1-z_1) - \rho_k z_1, \quad (83)$$

$$D_k = (1-\rho_k)z_1 + \rho_k(1-z_1) + \rho_k z_1. \quad (84)$$

Now we use the fact that each ρ_k converges to z_1 exponentially for each k. That is, we take each $\rho_k = z_1$ and substitute into the backward equation for T. The change of variables affects the derivatives, giving

$$\frac{\partial}{\partial \rho_k} \to \frac{kN_k}{N\mu_1}\frac{\partial}{\partial z_1}. \quad (85)$$

Now the backward equation is

$$-1 = \sum_k \left[v(z_1)\frac{kN_k}{N\mu_1}\frac{dT}{dz_1} + \frac{k^2 N_k}{2N^2\mu_1^2} D(z_1)\frac{d^2T}{dz_1^2} \right]. \quad (86)$$

Here

$$v(z_1) = -z_1^2, \quad (87)$$

$$D(z_1) = 2z_1(1-z_1) + z_1^2. \quad (88)$$

This simplifies to

$$-1 = v(z_1)\frac{dT}{dz_1} + \frac{\mu_2}{2N\mu_1^2} D(z_1)\frac{d^2T}{dz_1^2}. \quad (89)$$

Note that Eq. (89) has exactly the same form as Eq. (52). The primary difference is the appearance of μ_2/μ_1^2. To solve this, we follow the same procedure as the complete graph model and simply replace N with $N\frac{\mu_1^2}{\mu_2}$. This gives that the consensus time is

$$T \sim \sqrt{\frac{\pi^3}{8}\frac{N\mu_1^2}{\mu_2}}. \quad (90)$$

There are a few things to note about this solution. First, the form of the solution can be separated into the complete graph solution in Eq. (62) multiplied by the topological parameter $\sqrt{\mu_1^2/\mu_2}$. This may indicate that similar network models can be decomposed into a separable solution. That is, the consensus time for some models might be decomposed into the complete graph solution multiplied by an appropriate topological parameter. Second, the fact that $\mu_1^2/\mu_2 \leqslant 1$ indicates that sparse network topology only decreases the time to consensus. This implies that the complete graph solution is an upper bound for the consensus time for these models. Figure 4 depicts this solution for the complete graph and two scale-free networks.

B. Consensus model: Speaker first

We now consider a different model from the above listener-first model. Instead of assigning the first node to be the listener, we designate the first node to be the speaker. Then the speaker chooses to convey its message to a randomly chosen neighbor (the listener). We assume that only the listener updates their state based on this interaction. We also assume that consensus is an absorbing state, as in the listener-first case.

These assumptions produce a model with parameters α_{12}, β_{12}, β_{22}. All other parameters equal zero, and, for convenience, we drop the second subscript. We briefly outline the analysis, as it follows the same paradigm as the listener-first model. The transition probabilities for the model are given by

$$\Pr\{\Delta n_k = 1\} = \sum_i \sum_{j\in D_k} \frac{\alpha_1 A_{ij}}{Nk_i}\eta_i(1-\eta_j), \quad (91)$$

$$\Pr\{\Delta n_k = -1\} = \sum_i \sum_{j\in D_k} \frac{A_{ij}}{Nk_i}[\beta_1(1-\eta_i)\eta_j + \beta_2\eta_i\eta_j]. \quad (92)$$

We substitute $A_{ij} \to E[A_{ij}]$ and simplify the transition probabilities to obtain

$$p_k(\rho_k,\rho) = \alpha_1\frac{k}{\mu_1}\frac{N_k}{N}\rho(1-\rho_k), \quad (93)$$

$$q_k(\rho_k,\rho) = \frac{\beta_1 k}{\mu_1}\frac{N_k}{N}(1-\rho)\rho_k + \frac{\beta_2 k}{\mu_1}\frac{N_k}{N}\rho\rho_k. \quad (94)$$

We now generate the mean-field ODE system for $\bar{\rho}_k$ and obtain

$$\dot{\bar{\rho}}_k = \frac{k}{\mu_1}[\alpha_1\rho(1-\rho_k) - \beta_1(1-\rho)\rho_k - \beta_2\rho\rho_k]. \quad (95)$$

To simplify the analysis, we take $\alpha_1 = \beta_1 = \beta_2 = 1$. Note that the rate of change is proportional to k. We now multiply Eq. (95) by $\frac{N_k}{kN\mu_{-1}}$ and sum over k. Since $\alpha_1 = \beta_1 = 1$, this yields

$$\dot{z}_{-1} = -\frac{1}{\mu_1\mu_{-1}}\rho^2. \quad (96)$$

Since z_{-1} is monotonically decreasing, the system will globally converge to $\rho_k = 0$. This indicates that the convergence is very slow if the system is near consensus. This is due to the fact that the parameters are chosen at a phase transition, similarly to the complete graph. The slow convergence is also indicative of a stable center manifold. Linearizing the system for $\dot{\bar{\rho}}_k$ around $\rho_k = 0$ gives

$$\dot{\bar{\rho}}_k = \frac{k}{\mu_1}(\rho - \bar{\rho}_k). \quad (97)$$

This system has a Jacobian matrix given by

$$\frac{\partial \dot{\bar{\rho}}_k}{\partial \bar{\rho}_l} = \frac{kN_l}{N\mu_1} \qquad l \neq k$$
$$\frac{\partial \dot{\bar{\rho}}_k}{\partial \bar{\rho}_k} = \frac{k}{\mu_1}\left(\frac{N_k}{N} - 1\right). \tag{98}$$

It is a simple exercise to show that $\mathbf{1}$ is an eigenvector of the above Jacobian, whose eigenvalue is 0. In addition, all other eigenvalues have a negative real part by the Gershgorin circle theorem [47]. This indicates that the system tends to the center manifold near consensus, and the system then converges slowly to consensus. Furthermore, the eigenvector $\mathbf{1}$ indicates that each ρ_k tends to the same value. Making this substitution into Eq. (96) gives

$$\dot{z}_{-1} = -\frac{1}{\mu_1\mu_{-1}} z_{-1}^2, \tag{99}$$

which has the solution

$$z_{-1} = \left(C + \frac{t}{\mu_1\mu_{-1}}\right)^{-1} \tag{100}$$

for some constant C. Now we determine the consensus time by writing the backwards equation for T. The equation takes the form given in Eq. (82); however, v_k and D_k are given by

$$v_k = \frac{k}{\mu_1}[\rho(1-\rho_k) - (1-\rho)\rho_k - \rho\rho_k], \tag{101}$$
$$D_k = \frac{k}{\mu_1}[\rho(1-\rho_k) + (1-\rho)\rho_k + \rho\rho_k]. \tag{102}$$

We now apply the idea that each ρ_k tends to a common value given by the slow variable z_{-1}. We make the change of variables $\rho_k = z_{-1}$. The partial derivatives are now transformed by

$$\frac{\partial}{\partial \rho_k} \to \frac{k^{-1}N_k}{N\mu_{-1}}. \tag{103}$$

Substituting these into Eq. (82) yields the ODE for the consensus time given by

$$-\mu_1\mu_{-1} = v(z_{-1})\frac{dT}{dz_{-1}} + \frac{1}{2N}D(z_{-1})\frac{d^2T}{dz_{-1}^2}, \tag{104}$$

where

$$v(z_{-1}) = -z_{-1}^2, \tag{105}$$
$$D(z_{-1}) = 2z_{-1}(1-z_{-1}) + z_{-1}^2. \tag{106}$$

The reduced equation for T is remarkably similar to Eq. (52). Thus, we have translated the corresponding problem to a two-urn case with a single topological parameter $\mu_1\mu_{-1}$. Following the same steps as the complete graph case, we have shown that the consensus time for the speaker first model is

$$T \sim \mu_1\mu_{-1}\sqrt{\frac{\pi^3 N}{8}}. \tag{107}$$

It is simple to show that for every degree sequence, $\mu_1\mu_{-1} \geqslant 1$. This implies that this sparse graph structure will only slow down the rate of convergence, which is in contrast to the listener-first case which showed a decrease in consensus time. Since the listener-first model converges more quickly, it appears that imposing one's state on others is less effective in achieving consensus than listening to the messages of others. Also, the solution once again exhibits a separable form, in which a single topological parameter is multiplied by the complete graph solution.

VI. DISCUSSION

The two-urn models, and their extension to general network topology, are a rich class of models with a wide variety of behaviors. They are an extension of the Ehrenfest model of molecular diffusion, in which only one ball is chosen to change urn. By drawing two balls instead of one and allowing the redistribution to be stochastic, we introduce a class of models that have a wide range of applications from social dynamics to population genetics. Remarkably, for this class of models, we found that linearity of the ODE for mean density of urn A was necessary and sufficient for solvability by our generating function method. The seemingly unrelated notions of linearity of a differential equation and the diagonalization of a large transition matrix in a discrete Markov chain are, in fact, equivalent for these models. This method is a significant contribution to the field, since exact diagonalization of the Markov transition matrix for all N gives exact solutions to the probability distribution for all future times, consensus times, local times, etc. Furthermore, the eigenvectors describe the shape of the distribution as it tends toward stationary, and the eigenvalues describe the rate of convergence.

The work presented here on the two-urn models opens a wide range of questions for further study. In particular, we can contrive other random-walk models with more general polynomial transition probabilities, instead of restricting to the quadratics that correspond to specific urn-ball models. These in turn yield higher-order partial differential equations for a generating function G that in turn necessarily require new conditions for solvability. We have not shown that the two-urn models are the maximal set of models for high-degree transition probability distributions.

Another avenue that can be considered is to introduce multiple urns, yet retain a two ball interaction mechanism. This approach has been considered for the multistate voter model on the complete graph, which is a model with M urns [48]. A related model that has been analyzed by this procedure is the Naming Game with K opinions, which has $2^K - 1$ urns [17]. The work suggests that there is a great opportunity for studying highly complex models in these terms.

ACKNOWLEDGMENTS

This work was supported in part by the Army Research Office Grants No. W911NF-09-1-0254 and No. W911NF-12-1-467-0546. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government.

[1] P. Ehrenfest and T. Ehrenfest, Physik. Z. **8**, 311 (1907).
[2] T. M. Liggett, *Interacting Particle Systems* (Springer, Berlin, 2005).
[3] N. E. Henriksen and F. Y. Hansen, *Theories of Molecular Reaction Dynamics, The Microscopic Foundation of Chemical Kinetics* (Oxford University Press, Oxford, 2008).
[4] S. K. Upadhyay, *Chemical Kinetics and Reaction Dynamics* (Springer, New York, 2006).
[5] L. Frachebourg and P. L. Krapivsky, Phys. Rev. E. **53**, R3009(R) (1996).
[6] M. Kac, Am. Math. Mon. **54**, 369 (1947).
[7] P. A. P. Moran, Math. Proc. Cambr. Philos. Soc. **54**, 60 (1958).
[8] R. A. Blythe and A. J. McKane, J. Stat. Mech. (2007) P07018.
[9] T. Liggett, *Stochastic Interacting Systems: Contact, Voter, and Exclusion Processes* (Springer-Verlag, New York, 1999).
[10] P. Clifford and A. Sudbury, Biometrika **60**, 581 (1973).
[11] C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. **81**, 591 (2009).
[12] A. Baronchelli, V. Loreto, and L. Steels, Int. J. Mod. Phys. C. **19**, 785 (2008).
[13] J. Xie, S. Sreenivasan, G. Korniss, W. Zhang, C. Lim, and B. K. Szymanski, Phys. Rev. E. **84**, 011130 (2011).
[14] W. Zhang, C. Lim, S. Sreenivasan, J. Xie, B. Szymanski, and G. Korniss, Chaos **21**, 025115 (2011).
[15] W. Zhang, C. Lim, and B. K. Szymanski, Phys. Rev. E **86**, 061134 (2012).
[16] P. Sen and B. K. Chakrabarti, *Sociophysics, An Introduction* (Oxford University Press, Oxford, 2013).
[17] W. Pickering, B. K. Szymanski, and C. Lim, Phys. Rev. E **93**, 052311 (2016).
[18] F. Vazquez and V. M. Eguíluz, New J. Phys. **10**, 063011 (2008).
[19] V. Sood and S. Redner, Phys. Rev. Lett. **94**, 178701 (2005).
[20] W. Pickering and C. Lim, Phys. Rev. E **91**, 012812 (2015).
[21] C. Cooper, R. Elsässer, and T. Radzik, SIAM J. Discr. Math. **27**, 1748 (2013).
[22] S. Karlin and J. L. McGregor, Trans. Amer. Math. Soc. **85**, 489 (1957).
[23] E. S. Lee, Demography **3**, 47 (1966).
[24] J. J. Lewer and H. Van den Berg, Econ. Lett. **99**, 164 (2008).
[25] J. E. Anderson, Am. Econ. Rev. **69**, 106 (1979).
[26] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).
[27] M. E. J. Newman, Phys. Rev. E **66**, 016128 (2002).
[28] E. A. Bender and S. G. Williamson, *Foundations of Combinatorics with Applications* (Dover, New York, 2006).
[29] A. Edelman and G. Strang, Am. Math. Mon. **111**, 189 (2004).
[30] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1964).
[31] O. Krafft and M. Schaefer, J. Appl. Prob. **30**, 964 (1993).
[32] H. Dette, J. Appl Prob. **31**, 930 (1994).
[33] S. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*, Vol. 272 (Westview Press, Boulder, CO, 2001).
[34] W. Kornblum, *Sociology in a Changing World* (Cengage Learning, 2007).
[35] J. Berger and G. Le Mens, Proc. Natl. Acad. Sci. USA **106**, 8146 (2009).
[36] L. Dall'Asta, A. Baronchelli, A. Barrat, and V. Loreto, Phys. Rev. E **74**, 036105 (2006).
[37] A. Waagen, G. Verma, K. Chan, A. Swami, and R. D'Souza, Phys. Rev. E **91**, 022811 (2015).
[38] J. T. Cox, Ann. Probab. **17**, 1333 (1989).
[39] M. A. M. de Aguiar and Y. Bar-Yam, Phys. Rev. E **84**, 031901 (2011).
[40] R. Arratia, Ann. Probab. **9**, 909 (1981).
[41] J. Cardy and U. C. Täuber, Phys. Rev. Lett. **77**, 4780 (1996).
[42] T. Luczak, in *Proceedings of the Symposium on Random Graphs* (1989), pp. 165–182.
[43] M. Molloy and B. Reed, Random Struct. Algor. **6**, 161 (1995).
[44] M. Molloy and B. Reed, Combin. Probab. Comput. **7**, 295 (1998).
[45] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[46] R. Blythe, J. Phys. A **43**, 385003 (2010).
[47] S. Roman, *Advanced Linear Algebra*, 3rd ed. (Springer-Verlag, New York, 2008).
[48] W. Pickering and C. Lim, Phys. Rev. E **93**, 032318 (2016).