Community detectability and structural balance dynamics in signed networks

Megan Morrison [*]

Department of Applied Mathematics, University of Washington, Washington 98115, USA

Michael Gabbay [†]

Applied Physics Laboratory, University of Washington, Washington 98115, USA

We investigate signed networks with community structure with respect to their spectra and their evolution under a dynamical model of structural balance, a prominent theory of signed social networks. The spectrum of the adjacency matrix generated by a stochastic block model with two equal-size communities shows detectability transitions in which the community structure becomes manifest when its signal eigenvalue appears outside the main spectral band. The spectrum also exhibits "sociality" transitions involving the homogeneous structure representing the average tie value. We derive expressions for the eigenvalues associated with the community and homogeneous structure as well as the transition boundaries, all in good agreement with numerical results. Using the stochastically generated networks as initial conditions for a simple model of structural balance dynamics yields three outcome regimes: two hostile factions that correspond with the initial communities, two hostile factions uncorrelated with those communities, and a single harmonious faction of all nodes. The detectability transition predicts the boundary between the assortative and mixed two-faction states and the sociality transition predicts that between the mixed and harmonious states. Our results may yield insight into the dynamics of cooperation and conflict among actors with distinct social identities.

I. INTRODUCTION

Most research in network science has focused on networks that allow only positive ties. In signed networks, however, ties can take on negative values as well. In social systems, positive ties signify friendly or cooperative relationships between the individual or collective actors represented by the nodes whereas negative ties signify hostile or conflictual relationships between nodes. As examples, signed social networks have been used to represent interpersonal sentiments among students [1], supportive or critical references among opinion makers [2], relationships in online social networks [3], and alliances and military clashes among nations [4,5].

In this paper, we address community structure in signed networks and its implications for dynamics governed by structural balance, a theory commonly invoked in treatments of signed networks in social systems. In unsigned networks, community structure refers to the presence of clusters within networks characterized by relatively dense intracluster ties and sparse intercluster ties. A rich set of techniques have been developed to detect communities in unsigned networks [6]. Of particular relevance here, spectral analysis has proven to be a highly valuable tool for probing community structure [7,8]. For signed networks, the notion of community can be extended to accommodate negative ties by reversing the criteria for positive ties—there should be relatively sparse negative ties within communities and denser ties between them. At the present, however, the literature on community detection in signed networks is itself rather sparse in comparison with unsigned networks (e.g., Refs. [2,9–11]).

An important phenomenon of community structure in unsigned networks is that of community detectability [12–17]. Here, community structure can be present—in the sense that the tie-generating probabilities in a stochastic block model indeed favor ingroup over outgroup ties—but it is too weak to typically be discerned by analysis of the generated network. For large networks, a phase transition characterizes the passage from undetectable to detectable structure.

We show that detectability transitions also occur in signed networks. We generate our networks using a stochastic block model for two communities in an unweighted and undirected signed network (Sec. II). Examples of simulated networks with community structure that is detectable and undetectable are shown on the left in Figs. 1(a) and 1(b). We describe the transitions observed in the spectra of simulated networks in which outlying eigenvalues corresponding to meaningful signals merge with the main spectral band corresponding to noise (Sec. III). Two sets of spectral transitions are found: one corresponds to the detectability transition involving the two-community structure, while the other affects the ability to observe an overall tendency toward positive or negative tie formation, which we refer to as sociality transitions. Figure 1(c) shows an example of a network with a markedly positive average tie value generated in the regime in which an overall prosocial tendency can be reliably discerned.

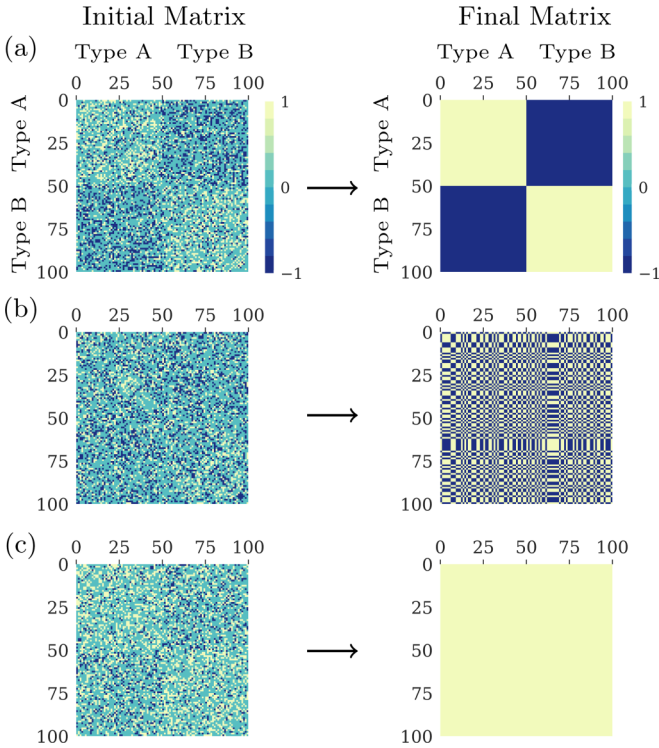[*]mmtree@uw.edu

[†]gabbay@uw.edu

FIG. 1. Evolution of networks with initial community structure under structural balance dynamics. (a) Moderate initial structuring by group identity leads to a completely connected network consisting of two factions sorted by identity. (b) Weak initial structure leads to two factions of mixed identities. (c) Strong initial positivity in the network leads to a single harmonious faction. Networks represented as adjacency matrices with ± 1, 0 tie values indicated by color. Initial networks generated by stochastic block model, Eqs. (2)–(4) with parameters $N = 100$ and $d_{\mathrm{in}}, d_{\mathrm{out}} = 0.4$ in all networks. $p_{\mathrm{out}}^+ = 0.3$ and $p_{\mathrm{in}}^+ = 0.7$ for (a), $p_{\mathrm{out}}^+ = 0.4$ and $p_{\mathrm{in}}^+ = 0.5$ for (b), and $p_{\mathrm{out}}^+ = 0.6$ and $p_{\mathrm{in}}^+ = 0.8$ for (c). Final networks represent the connectivity signs to which Eq. (1) converges (see Sec. VI).

We analytically calculate both the key eigenvalues and the transition conditions for large networks. In the main text, we use perturbation analysis to derive expressions for the signal eigenvalues (Sec. IV), which are then used to obtain the transition conditions by their equation with the main band edge eigenvalues (Sec. V), these edge eigenvalues being found using random matrix theory (Appendix A). We also present an alternative to our perturbation treatment that derives the signal eigenvalues on the basis of random matrix theory, in keeping with previous treatments of detectability (Appendix B) [13,14].

The spectral transitions have important implications for the outcomes of structural balance dynamics for networks possessing initial community structure. Structural balance theory, which postulates that triads with one or three negative edges will not endure, can be implemented as a deterministic, continuous time dynamical system (Sec. VI),

$$\frac{dY_{ij}}{dt} = \sum_{k=1}^{N} Y_{ik} Y_{kj}, \tag{1}$$

where t is time and N is the number of nodes [18,19]. This system evolves the connectivity Y_{ij} between nodes i and j as a function of their relationships with mutual neighbors: the product $Y_{ik} Y_{kj}$ increases their connectivity when they share a common inclination, positive or negative, toward k but decreases it if their inclinations are oppositely signed. This dynamic promotes balanced triads and eradicates unbalanced triads in the network. The model evolves into a fully connected network where either: (1) there are two hostile factions with only positive ties within each and only negative ties between them; or (2) all nodes are positively connected in a single harmonious faction. In either case, the final state is determined by the leading eigenvector of the initial network.

The driving role played by the leading eigenvector of the initial network in the structural balance evolution gives rise to a dynamical manifestation of the detectability transition when the leading eigenvector also carries the community structure signal. For the two-faction outcome, if the leading eigenvector corresponds to the two identity types in the stochastic block model, then the final factions will perfectly align with these identities as shown in Fig. 1(a). However, if the leading eigenvector is merely the edge of the main noise band, as occurs for weak initial structure below the detectability transition, then the composition of the final factions will not align with the identity types as seen in Fig. 1(b). An analogous transition to the single-faction outcome is generated by the sociality transition as seen in Fig. 1(c). Solutions of the structural balance model starting from networks randomly generated by the stochastic block model do indeed show sharp transitions between behavioral regimes whose boundaries agree with analytical predictions based on the detectability and sociality transitions (Sec. VII).

We discuss the potential implications of these results for conflict dynamics among actors with different identity types due to, for instance, ethnicity, religion, or ideology (Sec. VIII). In particular, conflicts such as civil wars may take on a binary nature. If the system starts out with weak identity-driven structure, then it will not be expected to polarize on the basis of identity. But complete identity polarization results even when initial affinities and animosities between identity types are fairly mild and even though identity itself plays no role in the micro-level conflict dynamics.

II. GENERATING AND REPRESENTING COMMUNITY STRUCTURE

Communities in an unsigned network are characterized by relatively dense within-community ties and sparse ties between communities. Community detection algorithms seek to discover these communities given an observed network [6,7,20]. Stochastic block models, which generate random networks with community structure by setting tie probabilities within and between blocks of nodes, have been used to investigate the behavior of community detection algorithms [21]. In this section, we describe the stochastic block model we use to generate our signed networks, the characterization of community structure via assortativity, and decomposition of the generated networks in terms of the eigenvectors of the average adjacency matrix and a random matrix.

A. Stochastic block model

Our construction starts with an undirected network of N nodes consisting of two identity groups A and B of equal size $N/2$, where $N \gg 1$. The A group nodes are indexed from 1 to $N/2$ and the B group from $N/2 + 1$ to N. **A** is the signed adjacency matrix where A_{ij} is the tie value between node i and node j, which can take on values of $\{1, -1, 0\}$ with 0 signifying the absence of a tie. As the network is undirected, the adjacency matrix is symmetric, $A_{ij} = A_{ji}$. The probability that a tie, positive or negative, will form between any given ingroup (A with A, B with B) node pair is d_{in}. Similarly, the tie formation probability between outgroup (A with B) node pairs is d_{out}. These tie formation probabilities are equivalent to the expected ingroup and outgroup tie densities and their average yields the expected tie density for the total network, $d = (d_{\text{in}} + d_{\text{out}})/2$. Given the presence of a tie between ingroup members, the conditional probability that it is positive is p_{in}^+ and that it is negative is $p_{\text{in}}^- = 1 - p_{\text{in}}^+$. Similarly, the positive and negative tie conditional probabilities between outgroup nodes are written p_{out}^+ and $p_{\text{out}}^- = 1 - p_{\text{out}}^+$. For brevity, we refer to p_{in}^+ and p_{out}^+ as the ingroup and outgroup *affinities* and p_{in}^- and p_{out}^- as the in and outgroup *animosities*.

The adjacency matrix can be written in terms of the following block structure:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{AA} & \mathbf{A}_{AB} \\ \mathbf{A}_{BA} & \mathbf{A}_{BB} \end{bmatrix}, \tag{2}$$

where each block is a random $N/2 \times N/2$ matrix. The diagonal blocks represent AA or BB ties, whose elements are set using the following probability distribution for the ingroup random variable A_{in}:

$$\mathbb{P}(A_{\text{in}} = k) = \begin{cases} d_{\text{in}} p_{\text{in}}^+, & k = 1, \\ d_{\text{in}}(1 - p_{\text{in}}^+), & k = -1, \\ 1 - d_{\text{in}}, & k = 0. \end{cases} \tag{3}$$

Since **A** is symmetric, there are $(N/2)(N/2 + 1)$ independent, identically distributed (i.i.d.) ingroup ties.

The off-diagonal blocks, corresponding to AB or BA ties, are transposes of each other resulting in $N^2/4$ i.i.d. outgroup ties, which are drawn according to the random variable A_{out}:

$$\mathbb{P}(A_{\text{out}} = k) = \begin{cases} d_{\text{out}} p_{\text{out}}^+, & k = 1, \\ d_{\text{out}}(1 - p_{\text{out}}^+), & k = -1, \\ 1 - d_{\text{out}}, & k = 0. \end{cases} \tag{4}$$

Equations (2)–(4) define the stochastic block model used to generate matrices with more or less community structure as seen on the left in Fig. 1 and for all the numerically generated spectra shown in this paper. Note that this model allows nonzero self-ties, unlike in many empirical networks, but this is a standard approximation that facilitates analytical treatment [13,14,22]. For large networks, the contribution of the N diagonal elements is negligible in comparison with that of the order N^2 off-diagonal elements. As we note below, the effect of removing self-ties on the average of **A** is to shift the eigenvalues by a constant that is independent of N. In addition, the model of structural balance dynamics, Eq. (1), allows for self-ties.
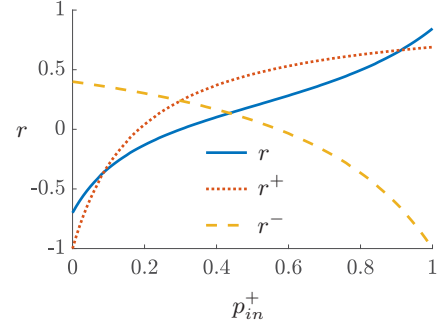

FIG. 2. Example of assortativity coefficients as a function of ingroup affinity. Values calculated using expected tie numbers. Parameters are $d_{\text{in}} = 0.5$, $d_{\text{out}} = 0.3$, $p_{\text{out}}^- = 0.7$, and $N = 100$.

Networks that are too sparse become disconnected and the structural balance dynamics of the isolated subgroups will evolve independently of each other as opposed to the holistic evolution we see when the network forms one connected graph. An Erdős-Rényi graph will very likely form a single connected graph if $p > \ln(N)/N$ [17,23]. Similarly, our stochastic block model matrices with two communities and density probabilities $d_{\text{in}} = a/N$ and $d_{\text{out}} = b/N$ will have a single giant component if $(a + b)/2 > 1$, and will very likely form a single connected graph if $d_{\text{in}} = a \ln(N)/N$, $d_{\text{out}} = b \ln(N)/N$ and $(a + b)/2 > 1$ [17].

B. Assortativity

Assortativity refers to the tendency for nodes of the same type to be more strongly connected than nodes of different types. We extend the standard definition of the assortativity coefficient for discrete node types [24] to our signed network case by calculating separate coefficients for the positive and negative tie networks and then essentially differencing them. We will use the signed network assortativity coefficient to characterize the regimes of the structural balance dynamics in Sec. VII.

First, considering the adjacency matrix of positive ties only, we let e_{ij}^+ denote the fraction of all positive ties that connect a node of type i to one of type j where $i, j \in \{A, B\}$. The assortativity coefficient r^+ for the network of positive ties, whose adjacency matrix elements are 1 if $A_{ij} > 0$ and zero otherwise, is then

$$r^+ = \frac{\sum_i e_{ii}^+ - \sum_i (a_i^+)^2}{1 - \sum_i (a_i^+)^2}, \tag{5}$$

where $a_i^+ = \sum_j e_{ij}^+$. The assortativity coefficient can range between -1 and 1. A network containing only ingroup (AA or BB) positive ties with no outgroup (AB, BA) ties is completely assortative, $r^+ = 1$, which in the social network context implies that people only cooperate with members of the same group. A network containing only outgroup ties is completely disassortative, $r^+ = -1$, implying cooperation across the two groups but not within them. We see this state in the example shown in Fig. 2 in which $r^+ = -1$ when the ingroup affinity $p_{\text{in}}^+ = 0$. As the ingroup affinity increases, r^+ increases but does not reach one as there are still outgroup ties due to the nonzero value of the fixed outgroup affinity.

The assortativity for the network of negative ties, r^-, is defined analogously to Eq. (5). Whereas positive values of r^+ imply that ingroup relations are more friendly than outgroup relations, assortative mixing in the network of negative ties implies more hostility within groups than between them. Thus, in Fig. 2 we see that $r^- > 0$ when there is complete ingroup animosity ($p_{\text{in}}^- = 1$ corresponding to $p_{\text{in}}^+ = 0$) and $r^- = -1$ when there is no ingroup animosity and so no negative ties within groups.

Accordingly, as we want positive values of our overall signed network assortativity coefficient r to signify that ingroup interactions tend to be more amicable than outgroup ones, we average r^+ and $-r^-$, yielding

$$r = \frac{r^+ - r^-}{2}, \tag{6}$$

which can take on values between -1 and 1. Figure 2 shows that r is negative for low ingroup affinity and positive for high ingroup affinity.

C. Adjacency matrix decomposition

This section presents a decomposition of the adjacency matrices generated by the stochastic block model into: (i) a signal component that results from the expected tie values generated by the ingroup and outgroup random variables, A_{in} and A_{out}; and (ii) a noise component due to random deviations from the expected values. This decomposition will form the starting point for our calculation of the network eigenvalues in Sec. IV. We write $\mathbf{A}$ as the sum of the average matrix $\langle \mathbf{A} \rangle$ and a random deviation matrix $\mathbf{X}$:

$$\mathbf{A} = \langle \mathbf{A} \rangle + \mathbf{X}. \tag{7}$$

Given the block structure of Eq. (2), $\langle \mathbf{A} \rangle$ can be written as

$$\langle \mathbf{A} \rangle = \begin{bmatrix} \langle \mathbf{A}_{AA} \rangle & \langle \mathbf{A}_{AB} \rangle \\ \langle \mathbf{A}_{BA} \rangle & \langle \mathbf{A}_{BB} \rangle \end{bmatrix}, \tag{8}$$

where each element of $\langle \mathbf{A}_{AA} \rangle$ and $\langle \mathbf{A}_{BB} \rangle$ is equal to $\langle A_{\text{in}} \rangle$ and each element of $\langle \mathbf{A}_{AB} \rangle$ and $\langle \mathbf{A}_{BA} \rangle$ is equal to $\langle A_{\text{out}} \rangle$. From Eqs. (3) and (4), we have

$$\langle A_{\text{in}} \rangle = d_{\text{in}}(2p_{\text{in}}^+ - 1), \tag{9}$$

$$\langle A_{\text{out}} \rangle = d_{\text{out}}(2p_{\text{out}}^+ - 1). \tag{10}$$

We define a couple of useful linear combinations of the in and outgroup expected tie values. We denote by μ the average over all the elements in $\langle \mathbf{A} \rangle$,

$$\mu = \frac{\langle A_{\text{in}} \rangle + \langle A_{\text{out}} \rangle}{2}, \tag{11}$$

and we denote by ν the half-difference between the in and outgroup expected tie values,

$$\nu = \frac{\langle A_{\text{in}} \rangle - \langle A_{\text{out}} \rangle}{2}. \tag{12}$$

Both μ and ν range from -1 to 1. Noting that $\langle A_{\text{in}} \rangle = \mu + \nu$ and $\langle A_{\text{out}} \rangle = \mu - \nu$, these expressions allow us to express $\langle \mathbf{A} \rangle$ as a sum of two outer products,

$$\langle \mathbf{A} \rangle = \mu N \mathbf{u}_H \mathbf{u}_H^T + \nu N \mathbf{u}_C \mathbf{u}_C^T, \tag{13}$$

where $\mathbf{u}_H$ and $\mathbf{u}_C$ are orthonormal N-dimensional vectors: $\mathbf{u}_H = \frac{1}{\sqrt{N}}[1, 1, ..., 1]^T$ and $\mathbf{u}_C = \frac{1}{\sqrt{N}}[1, ..., 1, -1, ..., -1]^T$, where the -1's align with the B block node indices. In fact, $\mathbf{u}_H$ and $\mathbf{u}_C$ are readily seen to be the two eigenvectors of $\langle \mathbf{A} \rangle$ with respective eigenvalues μN and νN:

$$\langle \mathbf{A} \rangle \mathbf{u}_H = \mu N \mathbf{u}_H, \tag{14}$$

$$\langle \mathbf{A} \rangle \mathbf{u}_C = \nu N \mathbf{u}_C. \tag{15}$$

The term containing $\mathbf{u}_H$ in Eq. (13) generates a homogeneous $N \times N$ matrix whose elements are all equal to μ, the global average tie value. Hence, we refer to $\mathbf{u}_H$ as the *homogeneous* eigenvector. The term containing $\mathbf{u}_C$ generates a matrix whose diagonal block elements are all equal to ν and whose off-diagonal block elements are $-\nu$ and so corresponds to the structure of ingroup and outgroup tie differences. Accordingly, $\mathbf{u}_C$ generates the community structure and we refer to it as the *contrast* eigenvector.

The homogeneous and contrast eigenvectors are *signal* eigenvectors whose ability to be distinguished from the noise generated by $\mathbf{X}$ has important implications for community detectability and structural balance dynamics. From this perspective, μ and ν can be regarded as natural parameters for the signal structure in the network and could be used in place of two of the parameters in the stochastic block model, for instance, the ingroup and outgroup affinities. Doing so is less intuitive from a simulation viewpoint, however.

While $\langle \mathbf{A} \rangle$ is rank 2, in general, if either μ or ν equals zero, then $\langle \mathbf{A} \rangle$ becomes a rank 1 matrix composed of either the homogeneous eigenvector or the contrast eigenvector. If $p_{\text{out}}^+ = 1/2$, then $\langle A_{\text{out}} \rangle = 0$ and $\mu = \nu$ and so the two eigenvalues of $\langle \mathbf{A} \rangle$ are degenerate, as is the case when $d_{\text{out}} = 0$ and the two blocks are disconnected from each other. However, taking $d_{\text{in}} = 0$ yields a bipartite network between the A and B blocks in which $\mu = \langle A_{\text{out}} \rangle / 2 = -\nu$, and so the homogeneous and contrast eigenvalues are equal and opposite. If μ and ν are both zero, then $\langle \mathbf{A} \rangle$ vanishes and $\mathbf{A}$ reduces to the noise matrix $\mathbf{X}$.

To remove self-ties from $\langle \mathbf{A} \rangle$, one can subtract $\langle A_{\text{in}} \rangle \mathbf{I}$ from Eq. (13), where $\mathbf{I}$ is the identity matrix. This shifts the signal eigenvalues by $-\langle A_{\text{in}} \rangle = -(\mu + \nu)$.

The noise matrix $\mathbf{X}$ is a symmetric matrix that can be written in the block form,

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{AA} & \mathbf{X}_{AB} \\ \mathbf{X}_{BA} & \mathbf{X}_{BB} \end{bmatrix}. \tag{16}$$

Since $\mathbf{X} = \mathbf{A} - \langle \mathbf{A} \rangle$, the elements of the ingroup blocks $\mathbf{X}_{AA}$ and $\mathbf{X}_{BB}$ can assume values in $\{1 - \langle A_{\text{in}} \rangle, -1 - \langle A_{\text{in}} \rangle, -\langle A_{\text{in}} \rangle\}$ that are distributed according to the random variable X_{in},

$$\mathbb{P}(X_{\text{in}} = k) = \begin{cases} d_{\text{in}} p_{\text{in}}^+, & k = 1 - \langle A_{\text{in}} \rangle, \\ d_{\text{in}}(1 - p_{\text{in}}^+), & k = -1 - \langle A_{\text{in}} \rangle, \\ 1 - d_{\text{in}}, & k = -\langle A_{\text{in}} \rangle. \end{cases} \tag{17}$$

Likewise, the entries of the outgroup blocks $\mathbf{X}_{AB} = \mathbf{X}_{BA}^T$ are distributed like X_{out},

$$\mathbb{P}(X_{\text{out}} = k) = \begin{cases} d_{\text{out}} p_{\text{out}}^+, & k = 1 - \langle A_{\text{out}} \rangle, \\ d_{\text{out}}(1 - p_{\text{out}}^+), & k = -1 - \langle A_{\text{out}} \rangle, \\ 1 - d_{\text{out}}, & k = -\langle A_{\text{out}} \rangle. \end{cases} \tag{18}$$

All the elements of $\mathbf{X}$ have zero mean as $\langle X_{\text{in}} \rangle = \langle X_{\text{out}} \rangle = 0$. The variances of X_{in} and X_{out} are given by $\sigma_{\text{in}}^2 = \langle X_{\text{in}}^2 \rangle$ and $\sigma_{\text{out}}^2 = \langle X_{\text{out}}^2 \rangle$, which are written in terms of the stochastic block model parameters as

$$\sigma_{\text{in}}^2 = d_{\text{in}} - d_{\text{in}}^2 (2 p_{\text{in}}^+ - 1)^2, \tag{19}$$

$$\sigma_{\text{out}}^2 = d_{\text{out}} - d_{\text{out}}^2 (2 p_{\text{out}}^+ - 1)^2. \tag{20}$$

These variances will appear as their average,

$$\sigma^2 = \frac{\sigma_{\text{in}}^2 + \sigma_{\text{out}}^2}{2}, \tag{21}$$

in the noise-induced correction to the signal eigenvalues calculated below. The average variance can also be related to the parameters μ and ν as follows:

$$\sigma^2 = \frac{d_{\text{in}} + d_{\text{out}}}{2} - \frac{d_{\text{in}}^2 (2 p_{\text{in}}^+ - 1)^2 + d_{\text{out}}^2 (2 p_{\text{out}}^+ - 1)^2}{2}, \tag{22}$$

$$= \frac{d_{\text{in}} + d_{\text{out}}}{2} - \frac{1}{2} (\langle A_{\text{in}} \rangle^2 + \langle A_{\text{out}} \rangle^2), \tag{23}$$

$$= \frac{d_{\text{in}} + d_{\text{out}}}{2} - \mu^2 - \nu^2, \tag{24}$$

where we have used Eqs. (9) and (10) in the second line and $\langle A_{\text{in}} \rangle^2 + \langle A_{\text{out}} \rangle^2 = (\mu + \nu)^2 + (\mu - \nu)^2 = 2(\mu^2 + \nu^2)$ in the third.

III. SIGNAL EIGENVALUE TRANSITIONS

Spectral analysis has been used to address the number and detectability of communities in unsigned networks by considering the leading eigenvalues that reside outside the (approximately) continuous main spectral band due to its generation as a random graph [8,13,14,25]. For undirected networks, the adjacency matrix is symmetric and hence has a real spectrum. The number of detectable communities is equivalent to the number of positive eigenvalues that lie beyond the main spectral band. Nadakuditi and Newman [13] showed the existence of, and analytically calculated, a detectability transition in which the community structure, as generated by a stochastic block model with two communities, while still present becomes no longer detectable. Under assortative tie formation, this transition occurs once the second eigenvalue of the adjacency matrix, which carries the community structure information, merges with the main spectral band. Using random matrix theory, the authors derived expressions for both leading eigenvalues and the edge of the spectral band, thereby enabling the analytical determination of the transition dependence.

Similar to unsigned networks, Fig. 3 illustrates that the spectra of our signed networks consist of a continuous band of eigenvalues originating from $\mathbf{X}$, the variability or noise in the system, and signal eigenvalues originating from $\langle \mathbf{A} \rangle$, the structure in the system. The edges of the main spectral band, $\pm \gamma$, are derived in Appendix A,

$$\gamma = 2 \sigma \sqrt{N}, \tag{25}$$

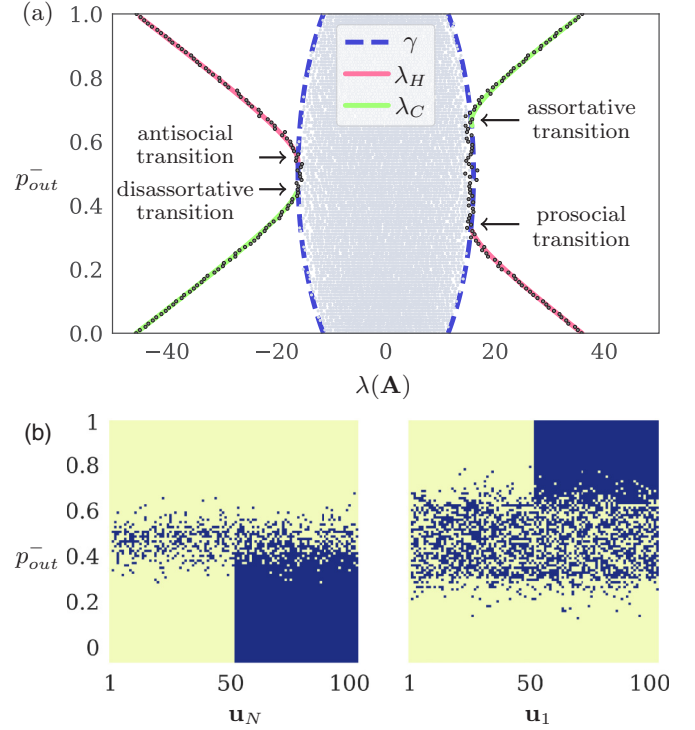
FIG. 3. (a) Eigenvalue spectrum of $\mathbf{A}$ as a function of outgroup animosity p_{out}^-. The other parameters remain constant: $N = 100$, $p_{\text{in}}^+ = 0.4$, $d_{\text{in}} = 0.5$, $d_{\text{out}} = 0.8$. (b) Signs (yellow, positive; blue, negative) of the components of the last, $\mathbf{u}_N$, and first, $\mathbf{u}_1$, eigenvectors as functions of p_{out}^-. The theoretical curves for λ_C (solid green), λ_H (solid pink), and γ (dashed blue) are calculated from Eqs. (26), (27), and (A15), respectively.

while the average contrast and homogeneous signal eigenvalues, $\langle \lambda_C \rangle$ and $\langle \lambda_H \rangle$, are derived in Sec. IV and Appendix B,

$$\langle \lambda_C \rangle = \nu N + \frac{\sigma^2}{\nu}, \quad |\nu| \geqslant \frac{\sigma}{\sqrt{N}}, \tag{26}$$

$$\langle \lambda_H \rangle = \mu N + \frac{\sigma^2}{\mu}, \quad |\mu| \geqslant \frac{\sigma}{\sqrt{N}}. \tag{27}$$

The formulas for $\langle \lambda_C \rangle$ and $\langle \lambda_H \rangle$ consist of their respective eigenvalues from $\langle \mathbf{A} \rangle$ and a correction proportional to the average variance. The conditions $|\nu|, |\mu| \geqslant \sigma/\sqrt{N}$ imply that $|\langle \lambda_C \rangle|, |\langle \lambda_H \rangle| \geqslant |\gamma|$ so that Eqs. (26) and (27) are only valid when outside of the main spectral band. Note that because of the self-averaging behavior for large N, we can effectively drop the expectation brackets and take $\lambda_C \approx \langle \lambda_C \rangle$ and $\lambda_H \approx \langle \lambda_H \rangle$.

The successive horizontal slices in Fig. 3(a) correspond to the eigenvalues of single instances of $\mathbf{A}$ generated by the stochastic block model as the outgroup animosity is increased. There are four points at which the outlying eigenvalues merge with the main band. Considering first the upper right of the plot, the largest eigenvalue, λ_1, is observed to detach from the main band for p_{out}^- greater than about 0.7. The right plot in Fig. 3(b), which depicts the signs of components of the first eigenvector $\mathbf{u}_1$, shows that $\mathbf{u}_1$ displays a two block structure for high p_{out}^-. Consequently, in this regime, the leading eigenvector corresponds to a perturbed version

of the contrast eigenvector, $\mathbf{u}_C$, of $\langle \mathbf{A} \rangle$. The point at which λ_1 emerges from the main band is then identified with the community detectability transition. Indeed, below this point $\mathbf{u}_1$ loses its block structure and rapidly takes on the appearance of random noise. We observe that in contrast to the analogous case in unsigned networks where the detectability transition involves the second eigenvalue of the adjacency matrix, here the two communities are no longer discernible when the leading eigenvalue merges with the main spectral band. We refer to the transition involving the merging of the contrast eigenvalue with the positive edge of the main band as the *assortative* transition because the communities are preferentially grouped by identity type. With an eye toward its dynamical significance when the contrast eigenvalue is leading, the two final factions produced by the structural balance dynamics are polarized by identity type after λ_1 emerges from the band.

The leading eigenvalue is observed to undergo another transition at $p_{out}^- \approx 0.3$. For lower outgroup animosities, $\mathbf{u}_1$ takes on a single block structure and so can be taken to result from a perturbation of the homogeneous eigenvector $\mathbf{u}_H$. This homogeneous structure disappears from $\mathbf{u}_1$ for p_{out}^- values above the transition. As the homogeneous eigenvalue carries information about the average tie value over all nodes, its emergence from the noise band can be considered a *sociality* transition. In particular, we refer to transitions that occur on the positive side of the noise band as *prosocial* transitions, in which a pattern of overall positive ties between nodes becomes apparent. The prosocial transition induces a transition in the structural balance dynamics from the two-faction equilibrium (not sorted by identity) to a single harmonious faction consisting of all nodes.

The lower left section of Fig. 3(a) shows the intersection of the last and least eigenvalue, λ_N, with the noise band at $p_{out}^- \approx 0.4$. For outgroup animosity values beneath this intersection, the last eigenvector $\mathbf{u}_N$ displays a two block structure as seen on the left plot of Fig. 3(b). However, although these blocks align with the A and B identity groups, the warmer outgroup than ingroup relations implies that the blocks are really disassortative "anti"-communities rather than assortative communities (prominent negative eigenvalues are also associated with disassortativity in unsigned networks [7]). Since this *disassortative* transition involves the least eigenvector, it has no significance with respect to the outcomes of the structural balance dynamics.

Finally, the other transition involving λ_N, seen in the upper left of Fig. 3(a), represents the emergence of the homogeneous eigenvalue and its corresponding single block structure from the noise band. It is a sociality transition and, in particular, an *antisocial* transition as it occurs on the negative side of the noise band signifying a conflictual relationship among nodes on average. The antisocial transition has no dynamical significance with respect to the structural balance dynamics.

Signal eigenvalues can occur on alternative sides of the spectral band as in Fig. 3(a) or on the same side as seen in Fig. 4(a). When λ_H and λ_C are on the same side of the spectral band, $p_{out}^- = 1/2$ is the point at which the signal eigenvalues cross one another since $\mu = \nu$. This affects which of the first two eigenvectors carries the community structure, but the structure itself remains apparent [Fig 4(b)]. However, the signal crossing does affect the balance dynamics,
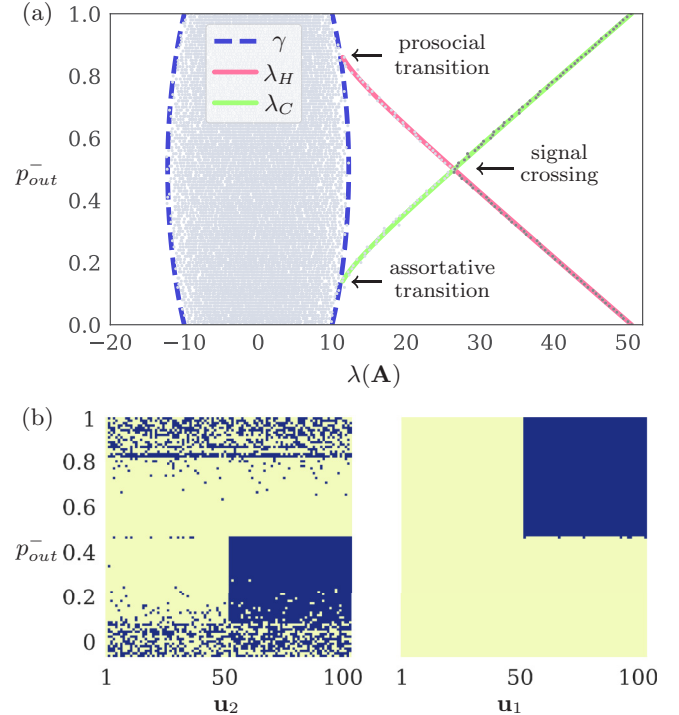

FIG. 4. (a) Spectrum of $\mathbf{A}$ as a function of p_{out}^- with constant parameters $N = 100$, d_{in}, $d_{out} = 0.5$, and $p_{in}^+ = 1$. (b) Signs (yellow positive, blue negative) of the components of the first and second eigenvectors, $\mathbf{u}_1$ and $\mathbf{u}_2$, as functions of p_{out}^-. The theoretical curves for λ_C (solid green), λ_H (solid pink), and γ (dashed blue) are calculated from Eqs. (26), (27), and (A15), respectively.

producing a transition between the harmonious and assortative outcomes.

IV. CALCULATION OF SIGNAL EIGENVALUES

In this section, we will derive formulas for the signal eigenvalues as a function of our stochastic block model parameters. We employ a perturbation treatment here but present an alternative derivation employing random matrix theory and complex analysis in Appendix B. Equation (7), which expresses the adjacency matrix $\mathbf{A}$ as the sum of its expected value $\langle \mathbf{A} \rangle$ and a matrix of random deviations $\mathbf{X}$, will form the starting point of our analysis. We consider $\langle \mathbf{A} \rangle$ as a given deterministic matrix with homogeneous and contrast eigenvalues and eigenvectors, Eqs. (14) and (15), that is subject to a perturbation from the independent noise matrix $\mathbf{X}$, which induces shifts to the signal eigenvalues and eigenvectors. A perturbation expansion to second order and the statistics of the noise matrix then yield the corrections to the signal eigenvalues. We note that in treating $\mathbf{X}$ as an independent perturbation to $\langle \mathbf{A} \rangle$, we temporarily suspend their linkage via the tie formation probabilities in the stochastic block model.

A. Perturbation expansion setup

We show the perturbation calculation for the case of the contrast eigenvalue. The homogeneous eigenvalue can be

obtained in precisely analogous fashion. The eigenvalue equation is

$$((\langle \mathbf{A} \rangle + \mathbf{X})\mathbf{v}_C = \lambda_C \mathbf{v}_C. \tag{28}$$

We write the perturbed eigenvector and eigenvalue up to second order as

$$\mathbf{v}_C = \mathbf{u}_C + \mathbf{u}^{(1)} + \mathbf{u}^{(2)}, \tag{29}$$

$$\lambda_C = \nu N + \lambda^{(1)} + \lambda^{(2)}, \tag{30}$$

where $\lambda^{(1)}, \lambda^{(2)}, \mathbf{u}^{(1)}, \mathbf{u}^{(2)}$ are the first- and second-order perturbations. As the unperturbed eigenvalue is $\mathcal{O}(\nu N)$, we divide Eq. (28) by N so that the zeroth-order equation is $\mathcal{O}(1)$. Then, using Eqs. (29) and (30), the eigenvalue equation becomes

$$\left(\frac{\langle \mathbf{A} \rangle}{N} + \frac{\mathbf{X}}{N} \right)(\mathbf{u}_C + \mathbf{u}^{(1)} + \mathbf{u}^{(2)})$$
$$= \left(\nu + \frac{\lambda^{(1)}}{N} + \frac{\lambda^{(2)}}{N} \right)(\mathbf{u}_C + \mathbf{u}^{(1)} + \mathbf{u}^{(2)}). \tag{31}$$

Before embarking upon our perturbation analysis, we specify the appropriate expansion parameter. To do so, we determine the orders of the $\langle \mathbf{A} \rangle$ and $\mathbf{X}$ matrices by evaluating their 2-norms. The 2-norm of each matrix is equivalent to its largest eigenvalue. Consequently, for the unperturbed matrix, $||\langle \mathbf{A} \rangle||_2$ is given by the larger of $N|\nu|$ or $N|\mu|$. Taking ν, μ to be $\mathcal{O}(1)$ therefore implies that $\langle \mathbf{A} \rangle$ is $\mathcal{O}(N)$. In Appendix A, using Wigner's semicircle law for the spectral density of a random matrix as well as matrix bounds, we determine that $||\mathbf{X}||_2$ is $\mathcal{O}(\sigma \sqrt{N})$. The ratio of the orders of $\mathbf{X}$ to $\langle \mathbf{A} \rangle$ is $\mathcal{O}(\sigma / \sqrt{N})$, and so successive orders in the perturbation series must diminish by a factor of $\sigma / \sqrt{N}$, which therefore serves as our expansion parameter. At a given N, the perturbation can be made arbitrarily small by letting σ go to zero. But in the large N regime, we need not constrain σ to be small.

Separating Eq. (31) out by expansion orders yields

$\mathcal{O}(1)$

$$\frac{\langle \mathbf{A} \rangle}{N} \mathbf{u}_C = \nu \mathbf{u}_C, \tag{32}$$

$\mathcal{O}(\sigma / \sqrt{N})$

$$\frac{\langle \mathbf{A} \rangle}{N} \mathbf{u}^{(1)} + \frac{\mathbf{X}}{N} \mathbf{u}_C = \nu \mathbf{u}^{(1)} + \frac{\lambda^{(1)}}{N} \mathbf{u}_C, \tag{33}$$

$\mathcal{O}(\sigma^2 / N)$

$$\frac{\langle \mathbf{A} \rangle}{N} \mathbf{u}^{(2)} + \frac{\mathbf{X}}{N} \mathbf{u}^{(1)} = \nu \mathbf{u}^{(2)} + \frac{\lambda^{(1)}}{N} \mathbf{u}^{(1)} + \frac{\lambda^{(2)}}{N} \mathbf{u}_C. \tag{34}$$

B. First-order treatment

To find the first-order eigenvalue perturbation, $\lambda^{(1)}$, we multiply both sides of Eq. (33) by $N\mathbf{u}_C^T$. Then using Eq. (13) and the orthonormality of $\mathbf{u}_C$ and $\mathbf{u}_H$ gives

$$\nu N \mathbf{u}_C^T \mathbf{u}^{(1)} + \mathbf{u}_C^T \mathbf{X} \mathbf{u}_C = \nu N \mathbf{u}_C^T \mathbf{u}^{(1)} + \lambda^{(1)}. \tag{35}$$

Solving for $\lambda^{(1)}$ yields

$$\lambda^{(1)} = \mathbf{u}_C^T \mathbf{X} \mathbf{u}_C \tag{36}$$
$$= \frac{1}{N} \left\{ \sum_{i,j=1}^{\frac{N}{2}} X_{ij} + \sum_{i,j=\frac{N}{2}+1}^{N} X_{ij} - 2 \sum_{i=1}^{\frac{N}{2}} \sum_{j=\frac{N}{2}+1}^{N} X_{ij} \right\}.$$

The first two terms in the braces above sum ties in the ingroup blocks *AA* and *BB*, respectively, each tie distributed as X_{in}, and the third term corresponds to the *AB* and BA outgroup ties, distributed as X_{out}. As $\lambda^{(1)}$ is equal to the sum of zero-mean random ingroup and outgroup variables, its mean therefore vanishes,

$$\langle \lambda^{(1)} \rangle = 0. \tag{37}$$

Turning to the variance, each element within the outgroup sum has variance σ_{out}^2, which becomes $4\sigma_{\text{out}}^2/N^2$ when the $2/N$ prefactor is included. The contribution to the variance from the $N^2/4$ outgroup variables is therefore σ_{out}^2. Similarly, for the ingroup sums, the symmetry of X implies that, neglecting the diagonal, there are approximately a total of $N^2/4$ independent variables each with variance $4\sigma_{\text{in}}^2/N^2$ so that the ingroup variance contribution is σ_{in}^2. Accordingly, the variance of $\lambda^{(1)}$ is

$$\text{Var}(\lambda^{(1)}) = \sigma_{\text{in}}^2 + \sigma_{\text{out}}^2 = 2\sigma^2. \tag{38}$$

To solve for $\mathbf{u}^{(1)}$, we write it as a vector decomposition and solve for the individual components,

$$\mathbf{u}^{(1)} = \mathbf{u}_{\|}^{(1)} + \mathbf{u}_{\perp}^{(1)}, \tag{39}$$

where $\mathbf{u}_{\|}^{(1)}$ is the component of $\mathbf{u}^{(1)}$ that is in the $\mathbf{u}_C, \mathbf{u}_H$ plane, and $\mathbf{u}_{\perp}^{(1)}$ is the component orthogonal to that plane. We find $\mathbf{u}_{\|}^{(1)}$ and $\mathbf{u}_{\perp}^{(1)}$ by multiplying both sides of the $\mathcal{O}(\sigma / \sqrt{N})$ equation, Eq. (33) by $\mathbf{u}_H^T$, the transpose of the homogeneous eigenvector of $\langle \mathbf{A} \rangle$, which gives

$$\mathbf{u}_H^T(\langle \mathbf{A} \rangle \mathbf{u}^{(1)} + \mathbf{X} \mathbf{u}_C) = \mathbf{u}_H^T(\nu N \mathbf{u}^{(1)} + \lambda^{(1)} \mathbf{u}_C), \tag{40}$$

which after employing the eigenvector properties becomes

$$\mu N \mathbf{u}_H^T \mathbf{u}^{(1)} + \mathbf{u}_H^T \mathbf{X} \mathbf{u}_C = \nu N \mathbf{u}_H^T \mathbf{u}^{(1)}. \tag{41}$$

Rearranging and noting that $\mathbf{u}_H^T \mathbf{u}^{(1)} = \mathbf{u}_H^T \mathbf{u}_{\|}^{(1)}$, we find

$$\mathbf{u}_H^T \mathbf{u}^{(1)} = \mathbf{u}_H^T \mathbf{u}_{\|}^{(1)} = \frac{\mathbf{u}_H^T \mathbf{X} \mathbf{u}_C}{\nu N - \mu N}, \tag{42}$$

which suggests the following solution for $\mathbf{u}_{\|}^{(1)}$,

$$\mathbf{u}_{\|}^{(1)} = \frac{[\mathbf{X} \mathbf{u}_C]_{\|}}{\nu N - \mu N}. \tag{43}$$

Writing $[\mathbf{X} \mathbf{u}_C]_{\|}$ as a decomposed projection onto $\mathbf{u}_C$ and $\mathbf{u}_H$,

$$[\mathbf{X} \mathbf{u}_C]_{\|} = (\mathbf{u}_C^T \mathbf{X} \mathbf{u}_C)\mathbf{u}_C + (\mathbf{u}_H^T \mathbf{X} \mathbf{u}_C)\mathbf{u}_H, \tag{44}$$

allows us to write $\mathbf{u}_{\|}^{(1)}$ as

$$\mathbf{u}_{\|}^{(1)} = \frac{(\mathbf{u}_C^T \mathbf{X} \mathbf{u}_C)}{\nu N - \mu N}\mathbf{u}_C + \frac{(\mathbf{u}_H^T \mathbf{X} \mathbf{u}_C)}{\nu N - \mu N}\mathbf{u}_H. \tag{45}$$

We now seek to solve for $\mathbf{u}_\perp^{(1)}$. Using the decomposition Eq. (39) in the first-order equation, Eq. (33), gives

$$\langle \mathbf{A} \rangle \big[\mathbf{u}_\shortparallel^{(1)} + \mathbf{u}_\perp^{(1)} \big] + \mathbf{X}\mathbf{u}_C = \nu N \big[\mathbf{u}_\shortparallel^{(1)} + \mathbf{u}_\perp^{(1)} \big] + \lambda^{(1)} \mathbf{u}_C. \quad (46)$$

Retaining only the terms that have components orthogonal to the $\mathbf{u}_C$, $\mathbf{u}_H$ plane and rearranging yields

$$\mathbf{u}_\perp^{(1)} = \frac{[\mathbf{X}\mathbf{u}_C]_\perp}{\nu N}. \quad (47)$$

Using the substitution $[\mathbf{X}\mathbf{u}_C]_\perp = \mathbf{X}\mathbf{u}_C - [\mathbf{X}\mathbf{u}_C]_\shortparallel$ then gives the solution for the orthogonal component

$$\mathbf{u}_\perp^{(1)} = \frac{\mathbf{X}\mathbf{u}_C - \big[(\mathbf{u}_C^T \mathbf{X}\mathbf{u}_C)\mathbf{u}_C + (\mathbf{u}_H^T \mathbf{X}\mathbf{u}_C)\mathbf{u}_H \big]}{\nu N}. \quad (48)$$

Combining Eqs. (45) and (48) gives the solution for the first-order perturbation to the eigenvector $\mathbf{u}^{(1)}$,

$$\mathbf{u}^{(1)} = \frac{\mathbf{X}\mathbf{u}_C}{\nu N} + \frac{\mu (\mathbf{u}_C^T \mathbf{X}\mathbf{u}_C)}{\nu N(\nu - \mu)}\mathbf{u}_C + \frac{\mu (\mathbf{u}_H^T \mathbf{X}\mathbf{u}_C)}{\nu N(\nu - \mu)}\mathbf{u}_H. \quad (49)$$

C. Second-order treatment

Having found the first-order eigenvalue and eigenvector perturbations, $\lambda^{(1)}$ and $\mathbf{u}^{(1)}$, we can now solve for the second-order correction $\lambda^{(2)}$. We multiply both sides of Eq. (34) by $\mathbf{u}_C^T$ and then solve to get

$$\lambda^{(2)} = \mathbf{u}_C^T \mathbf{X}\mathbf{u}^{(1)} - \lambda^{(1)}\mathbf{u}_C^T \mathbf{u}^{(1)} \quad (50)$$

$$= \frac{\mathbf{u}_C^T \mathbf{X}^2 \mathbf{u}_C}{\nu N} - \frac{(\mathbf{u}_C^T \mathbf{X}\mathbf{u}_C)^2}{\nu N} + \frac{\mu (\mathbf{u}_H^T \mathbf{X}\mathbf{u}_C)^2}{\nu N(\nu - \mu)}, \quad (51)$$

where Eqs. (36) and (49) have been used to obtain the second line.

We seek the expected value $\langle \lambda^{(2)} \rangle$ and consider the right-hand terms of Eq. (51) in succession. Expanding the expected value of the first term yields

$$\frac{\langle \mathbf{u}_C^T \mathbf{X}^2 \mathbf{u}_C \rangle}{\nu N} = \frac{1}{\nu N}\frac{1}{N}\left\{ \sum_{i,j=1}^{\frac{N}{2}}\langle (\mathbf{X}^2)_{ij} \rangle + \sum_{i,j=\frac{N}{2}+1}^{N}\langle (\mathbf{X}^2)_{ij} \rangle \right.$$
$$\left. - 2\sum_{i=1}^{\frac{N}{2}}\sum_{j=\frac{N}{2}+1}^{N}\langle (\mathbf{X}^2)_{ij} \rangle \right\}, \quad (52)$$

where $(\mathbf{X}^2)_{ij} = \sum_{k=1}^{N} X_{ik}X_{kj}$. As the elements of $\mathbf{X}$ are independent, the cross-element terms in this sum have vanishing expectation: $\langle X_{ik}X_{kj} \rangle = 0$ for $i \neq j$. When $i = j$, the value of $\langle X_{ik}^2 \rangle$ is either the ingroup or outgroup variance: $\langle X_{ik}^2 \rangle = \sigma_{\text{in}}^2$ if $i,k \leqslant N/2$ or $i,k > N/2$; $\langle X_{ik}^2 \rangle = \sigma_{\text{out}}^2$ otherwise. Accordingly, the expectations for the elements of $\mathbf{X}^2$ are given by

$$\langle (\mathbf{X}^2)_{ij} \rangle = \begin{cases} \frac{N}{2}\sigma_{\text{in}}^2 + \frac{N}{2}\sigma_{\text{out}}^2 = N\sigma^2, & i = j, \\ \\ 0, & i \neq j. \end{cases} \quad (53)$$

The above equation reduces the double sums in the first two terms in Eq. (52) to single sums over $\langle (\mathbf{X}^2)_{ii} \rangle = N\sigma^2$, which can then be combined. The last term, which contains only

off-diagonal elements of $\langle (\mathbf{X}^2) \rangle$, vanishes. The contribution of the first term in Eq. (51) to $\langle \lambda^{(2)} \rangle$ is therefore

$$\frac{\langle \mathbf{u}_C^T \mathbf{X}^2 \mathbf{u}_C \rangle}{\nu N} = \frac{1}{\nu N}\frac{1}{N}\sum_{i=1}^{N} N\sigma^2 \quad (54)$$

$$= \frac{\sigma^2}{\nu}. \quad (55)$$

We now turn to the second and third terms on the right-hand side of Eq. (51). The numerator of the second term, $(\mathbf{u}_C^T \mathbf{X}\mathbf{u}_C)^2$, involves the square of $\lambda^{(1)}$ by Eq. (36). Hence, its expected value is equivalent to the variance of $\lambda^{(1)}$ (which has zero mean), and so goes as σ^2 as shown above. Consequently, the expected value of the second term goes as σ^2/N. The same argument holds for the third term as its numerator depends on $\mathbf{u}_H^T \mathbf{X}\mathbf{u}_C$, which is likewise tantamount to the sum over the random ingroup and outgroup variables. Therefore, in the large-N regime of concern here, the second and third terms, which have a $1/N$ dependence, can be neglected in comparison with the first term, which is independent of N. Accordingly, Eq. (55) gives the second-order eigenvalue perturbation,

$$\langle \lambda^{(2)} \rangle = \frac{\sigma^2}{\nu}. \quad (56)$$

Having found that the first-order perturbation vanishes on average and given the second-order perturbation above, we arrive at the approximate solution for the expected value of the contrast eigenvalue of Eq. (28),

$$\langle \lambda_C \rangle = \nu N + \frac{\sigma^2}{\nu}. \quad (57)$$

Since the ratio of the second-order correction to the unperturbed eigenvalue goes as σ^2/N, the expansion parameter, given by its square root, is therefore $\mathcal{O}(\sigma/\sqrt{N})$ as stated at the beginning of this calculation.

A similar calculation, this time expanding about the homogeneous eigenvector $\mathbf{u}_H$, gives us the solution for λ_H, which simply involves swapping out ν for μ in the preceding equation,

$$\langle \lambda_H \rangle = \mu N + \frac{\sigma^2}{\mu}. \quad (58)$$

The analytical expressions for the signal eigenvalues, Eqs. (57) and (58), are plotted in Fig. 3(a) for values outside the main spectral band. They are observed to be in good agreement with the outlying eigenvalues of the numerical spectrum. This is the case even though, rather than an average over many generated networks, each horizontal slice represents just one instance, a reflection of the self-averaging behavior of large random networks. These expressions also work well in the nonsparse limit as shown in Fig. 5 for the case where the contrast eigenvalue becomes the leading eigenvalue beyond the assortative transition. The predicted λ_C is observed to separate from the spectral edge past a critical density, which decreases with network size, and shows good agreement with the first eigenvalue of the simulated network, particularly for the two larger networks. For the special case $\mu = \nu$, the same signal eigenvalue expressions still hold but we note how this case affects the derivation. When $\mu \neq \nu$, the last term on the right-hand side of Eq. (51) could be neglected above as
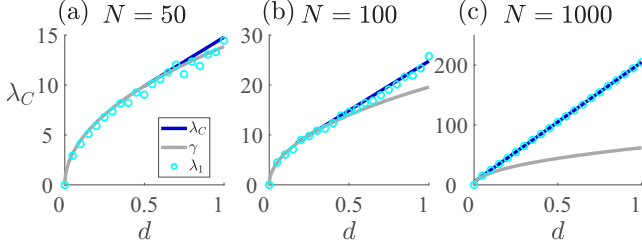
FIG. 5. Theoretical contrast [λ_C, Eq. (57)] and band edge [γ, Eq. (A15)] eigenvalues, along with the simulated leading eigenvalue (λ_1), as a function of network density. (a) $N = 50$, (b) $N = 100$, (c) $N = 1000$. λ_C is plotted for network density values where it meets or exceeds γ. One leading eigenvalue instance is computed for each d value. $d = d_{in} = d_{out}$, $p_{in}^+ = 0.6$, and $p_{out}^+ = 0.4$.

it is a factor of $1/N$ smaller than the first term. For $\mu = \nu$, however, the third term is singular. Yet, Eq. (41) implies that $\mathbf{u}_H^T \mathbf{X} \mathbf{u}_C = 0$ for $\mu = \nu$ (or is higher order for $\mu - \nu \sim 1/N$) and so the seemingly problematic third term does not arise. For the special cases $\nu = 0$ or $\mu = 0$, $\langle\lambda_C\rangle = 0$ or $\langle\lambda_H\rangle = 0$, respectively, and $\langle\mathbf{A}\rangle$ becomes rank 1. Figure 7(b) shows example spectra for $\mu = 0$ in which λ_C is the only signal eigenvalue.

The symmetric forms of the expressions for λ_C and λ_H reflect the fact that an orthogonal transformation $\mathbf{K}$ exists that transforms the contrast and homogeneous eigenvectors of $\langle\mathbf{A}\rangle$ into each other, that is, $\mathbf{u}_H' = \mathbf{K}\mathbf{u}_C$ and $\mathbf{u}_C' = \mathbf{K}\mathbf{u}_H$ where the primes denote the transformed system. Specifically, $\mathbf{K}$ is the diagonal matrix $\text{diag}(1, \ldots, 1, -1, \ldots, -1)$ where the negative values start at index $N/2 + 1$. It is its own inverse, $\mathbf{K}^{-1} = \mathbf{K}$. The transformation of the expected adjacency matrix, $\langle\mathbf{A}'\rangle = \mathbf{K}\langle\mathbf{A}\rangle\mathbf{K}$, flips the sign of the off-diagonal blocks so that $\langle A_{out}'\rangle = -\langle A_{out}\rangle$. Therefore, by Eqs. (11) and (12), $\mu' = \nu$ and $\nu' = \mu$, which swaps the perturbed signal eigenvalues, $\lambda_H' = \lambda_C$ and $\lambda_C' = \lambda_H$.

An alternative calculation of the signal eigenvalues based on random matrix theory and complex analysis is presented in the Appendix B. We find the same formulas for the signal eigenvalues [see Eq. (B29)] as have been derived here.

V. TRANSITION BOUNDARIES

In this section, we derive theoretical predictions for the boundaries of the detectability and sociality transitions. As discussed in Sec. III, these transitions occur when the signal eigenvalues merge with the main spectral band. From Eq. (A15), the edges of the main band of $\mathbf{X}$ are given by $\pm 2\sigma\sqrt{N}$, a formula that is a straightforward adaptation of the band edge of Wigner's semicircle distribution. We consider the community detectability transitions first, that is, those involving the contrast eigenvector $\mathbf{u}_C$, whose eigenvalue is given by Eq. (57). The detectability transition will therefore occur when

$$\nu N + \frac{\sigma^2}{\nu} = 2\sigma\sqrt{N}. \tag{59}$$

Solving for the critical value ν_* yields

$$\nu_* = \frac{\sigma(\nu_*)}{\sqrt{N}}. \tag{60}$$

The notation $\sigma(\nu_*)$ serves as a reminder that ν_* also appears on the righthand side due to the functional dependence of σ given by Eq. (24). The community structure is detectable when $|\nu| > \nu_*$. In particular, the assortative transition occurs for $\nu = \nu_*$ and the detectability transition for disassortative structure occurs for $\nu = -\nu_*$.

We observe that the transition condition Eq. (60) also results by setting the noise power equal to the signal power. Defining the noise power as the projection of $\mathbf{X}^2$ onto $\mathbf{u}_C$, its average, $\langle\mathbf{u}_C^T\mathbf{X}^2\mathbf{u}_C\rangle$, is found using Eq. (55) to be $N\sigma^2$. One could also arrive at this value by considering how much of the total noise variance, $N^2\sigma^2$, is carried on average by each of N randomly chosen orthogonal basis vectors. Equating the signal power to the average noise power, $\nu^2 N^2 = N\sigma^2$, yields Eq. (60).

The sociality transitions associated with the homogeneous signal occur when the homogeneous eigenvalue given by Eq. (58) equals the band edge eigenvalue. This yields a critical value μ_*,

$$\mu_* = \frac{\sigma(\mu_*)}{\sqrt{N}}. \tag{61}$$

The prosocial transition occurs for $\mu = \mu_*$ and the antisocial transition occurs for $\mu = -\mu_*$.

We now unpack the transition conditions derived above to express them in alternative ways in parameter space that will further intuitive understanding of the transition behavior and allow for connection with simulation results.

First, we substitute Eq. (24) for σ^2 in the detectability condition Eq. (60) to yield

$$N\nu^2 = \tfrac{1}{2}(d_{in} + d_{out}) - \mu^2 - \nu^2. \tag{62}$$

We point out that $(d_{in} + d_{out})/2$ is simply the overall tie density in the network. For a sparse network, $d_{in}, d_{out} \ll 1$, we can neglect the μ^2 and ν^2 terms on the right-hand side, so that the detectability transitions occur at

$$\pm\nu_* = \pm\sqrt{\frac{1}{2N}(d_{in} + d_{out})}. \tag{63}$$

The positive sign corresponds to the assortative transition and the negative sign corresponds to the disassortative transition. As ν can be regarded as a natural parameter for the community structure, this structure (assortative or disassortative) becomes easier to detect as the network becomes more sparse since ν_* shifts to smaller values (but care should be taken to distinguish the behavior of ν from that of the affinities and animosities, which can behave oppositely with density as in Eq. (72) below). Weaker structure is also more detectable as the size of the network grows, as was already apparent from Eq. (60).

We now substitute into Eq. (62) the definitions Eqs. (11) and (12) for μ and ν and rearrange to obtain

$$0 = \langle A_{out}\rangle^2 - \frac{2N}{N+2}\langle A_{out}\rangle\langle A_{in}\rangle$$
$$+ \langle A_{in}\rangle^2 - \frac{2}{N+2}(d_{in} + d_{out}), \tag{64}$$

which can be solved for the critical value of $\langle A_{\text{out}} \rangle$ (omitting the asterisk),

$$\langle A_{\text{out}} \rangle = \frac{N}{N+2} \langle A_{\text{in}} \rangle$$
$$\pm \sqrt{\frac{2}{N+2}(d_{\text{in}} + d_{\text{out}}) - \frac{4(N+1)}{(N+2)^2} \langle A_{\text{in}} \rangle^2}. \quad (65)$$

To write the detectability transitions completely in terms of the block model probabilities, we substitute Eqs. (9) and (10) for $\langle A_{\text{in}} \rangle$ and $\langle A_{\text{out}} \rangle$ into Eq. (65). Solving for the outgroup animosity and taking the large N limit yields

$$p_{\text{out}}^- = \frac{1}{2} - \frac{d_{\text{in}}}{d_{\text{out}}}\left(p_{\text{in}}^+ - \frac{1}{2}\right)$$
$$\pm \frac{1}{d_{\text{out}}}\sqrt{\frac{d_{\text{in}} + d_{\text{out}} - 8d_{\text{in}}^2\left(p_{\text{in}}^+ - \frac{1}{2}\right)^2}{2N}}, \quad (66)$$

where the positive sign corresponds to the assortative transition. Neglecting the second term inside the square root yields the sparse limit, equivalent to Eq. (63).

For the sociality transitions, the sparse limit results in the condition

$$\pm \mu_* = \pm\sqrt{\frac{1}{2N}(d_{\text{in}} + d_{\text{out}})}. \quad (67)$$

The positive and negative signs correspond to the prosocial and antisocial transitions, respectively. The critical value of the outgroup animosity is given by

$$p_{\text{out}}^- = \frac{1}{2} + \frac{d_{\text{in}}}{d_{\text{out}}}\left(p_{\text{in}}^+ - \frac{1}{2}\right)$$
$$\pm \frac{1}{d_{\text{out}}}\sqrt{\frac{d_{\text{in}} + d_{\text{out}} - 8d_{\text{in}}^2\left(p_{\text{in}}^+ - \frac{1}{2}\right)^2}{2N}}, \quad (68)$$

where the negative sign is used for the prosocial transition.

VI. STRUCTURAL BALANCE DYNAMICS

In its simplest incarnation, structural balance theory considers the stability of triads. Triads with all positive ties ("the friend of my friend is my friend") or two negative ties ("the enemy of my enemy is my friend") are considered balanced and so stable. In contrast, a triad with an odd number of negative ties will be unbalanced. For fully connected networks, assuming that all triads must be balanced over time implies that the system achieves either a state of global harmony in which all nodes are positively connected or two hostile camps with positive connections within each camp and negative connections between them [26]. Empirical signed networks in social systems such as international relations, student relationships, and online social networks have been found to be approximately balanced [1,3,27], exhibiting a tendency toward partition into two factions.

Although the concept of balance can be extended to arbitrary-length cycles, the triadic notion has motivated the construction of dynamical systems models that evolve the relationship between a pair of nodes as a function of their relationships with their network neighbors [18,19,28]. As

noted when Eq. (1) was introduced, if both members of a dyad have a positive relationship with a third node, then that will act toward making the focal dyad's relationship more positive. In contrast, having oppositely signed relationships with the third node will contribute a force pulling the dyad toward a more conflictual relationship. Unbalanced triads wither away under these dynamics. Building upon Ref. [18], Marvel *et al.* [19] demonstrated that the model of structural balance dynamics defined by Eq. (1) almost always achieves a balanced state starting from random initial conditions. Equation (1) can be written as a matrix equation,

$$\frac{d\mathbf{Y}}{dt} = \mathbf{Y}^2, \quad (69)$$

where $\mathbf{Y}$ is the matrix of signed and continuous connectivity values, Y_{ij}, between node pairs. In support of its empirical relevance, Ref. [19] found that when implemented upon the initial network of several real world systems, this model well predicts the observed final network.

Equation (69) is the matrix form of a Riccati equation and has the following closed form solution [19]:

$$\mathbf{Y}(t) = \mathbf{Y}(0)[\mathbf{I} - \mathbf{Y}(0)t]^{-1}. \quad (70)$$

The elements Y_{ij} diverge to positive or negative infinity in a finite time t_f and so the solution only holds for $t < t_f$.

For the purposes of analyzing community structure, we convert the connectivity matrix $\mathbf{Y}$ to an adjacency matrix $\mathbf{A}$ with discrete values ± 1 and 0 by taking the sign of the connectivity values so that $A_{ij} = \text{sgn}(Y_{ij})$. The leading eigenvector of the initial connectivity matrix, $\mathbf{Y}(0)$, grows fastest and so dominates the solution as $t \to t_f$. As a result, the final adjacency matrix $\mathbf{A}_f$ corresponding to $\mathbf{Y}$ as $t \to t_f$ converges to the outer product,

$$\mathbf{A}_f = \mathbf{u}_1 \mathbf{u}_1^T, \quad (71)$$

where $\mathbf{u}_1$ consists of the signs of the leading eigenvector of $\mathbf{Y}(0)$.

The rank 1 structure of $\mathbf{A}_f$ toward which the connectivity matrix converges implies that the final network must partition into either two hostile factions or one harmonious community as consistent with the expectations of structural balance theory [19]. The final network consists of a single harmonious faction if the components of $\mathbf{u}_1$ are of uniform sign, but consists of two hostile factions if $\mathbf{u}_1$ contains both positive and negative values [19]. Note that these results hold only if there is a single dominant eigenvalue and the graph is connected. When $d_{\text{out}} = 0$, the graph is disconnected and the isolated identity blocks will evolve independently of each other and the connectivity matrix will become rank 2 after structural balance dynamics. Although the first two eigenvalues of $\langle \mathbf{A} \rangle$ are equal in the $p_{\text{out}}^+ = 1/2$ case as well, the network remains connected and, due to stochasticity, one eigenvalue will inevitably be slightly larger in the realized $\mathbf{A}$, causing it to generate the ultimate rank 1 state. The case of $d_{\text{in}} = 0$ is also connected and so evolves to rank 1.

We will investigate the evolution of networks with community structure under the structural balance model (69). The initial connectivity matrix is taken to be proportional to an initial adjacency matrix $\mathbf{A}_0$ generated using the stochastic block model, in particular $\mathbf{Y}(0) = \mathbf{A}_0/N$. Figure 6(a) shows an
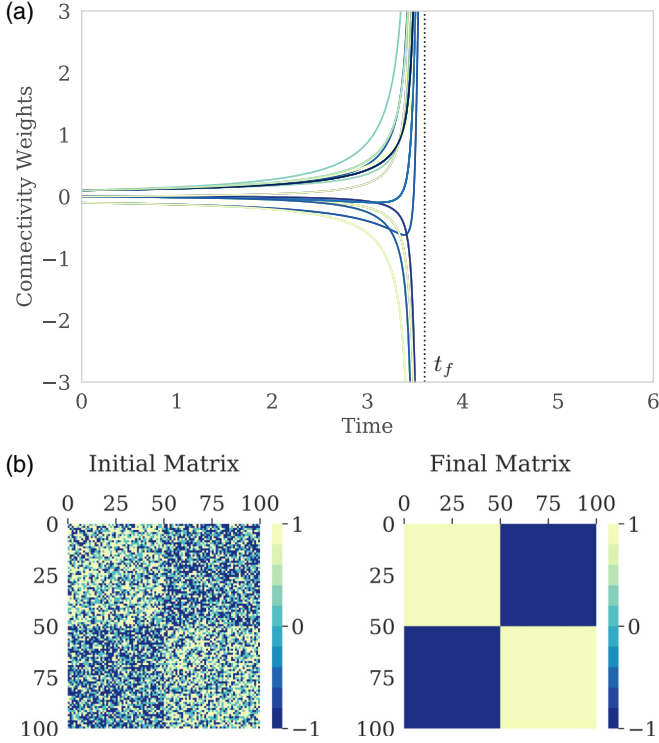
FIG. 6. Structural balance evolution of a network with community structure. (a) Connectivity network weights $Y_{ij}(t)$ over time evolved by Eq. (69) from an initial network generated by a stochastic block model (with values rescaled by $1/N$). (b) Initial and final adjacency matrices. Initial matrix parameters $N = 100$, $d_{in} = 0.7$, $d_{out} = 0.7$, $p_{in}^+ = 0.65$, and $p_{out}^+ = 0.35$.
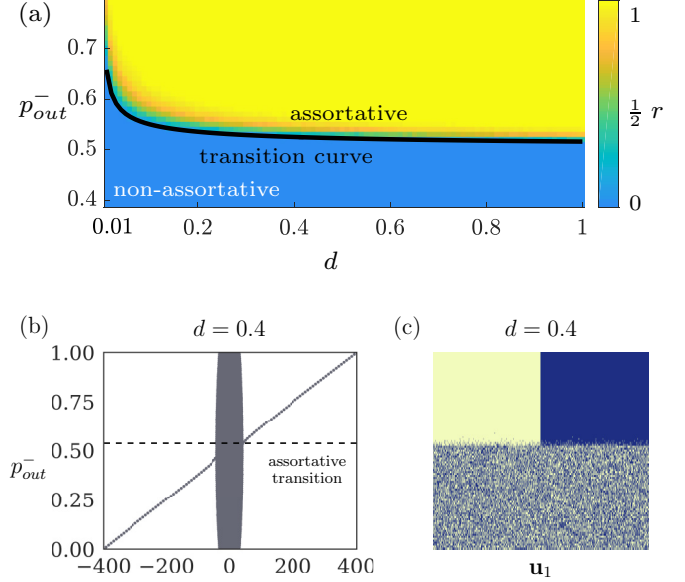

FIG. 7. (a) Assortativity of final adjacency matrix as a function of d and p_{out}^- for network size $N = 1000$ averaged over 400 simulations. Solid curve is the theoretical transition boundary given by Eq. (72). Note that the theoretical curve corresponds to the early part of the transition in which r just begins to rise, whereas the more visually distinctive yellow-cyan interface marks the middle of the transition. (b) Initial adjacency matrix spectra for $d = 0.4$ and increasing p_{out}^-. Dashed line indicates theoretical transition point. (c) Component signs for the leading eigenvector $\mathbf{u}_1$ of the initial adjacency matrix. The disassortative transition is not dynamically relevant so only the upper part of the p_{out}^- scale is plotted in (a).

example of the evolution of network connectivity values over time from a $\mathbf{Y}(0)$ corresponding to the $\mathbf{A}_0$ shown on the left in Fig. 6(b). We see that the $Y_{ij} \to \pm\infty$ and the final adjacency matrix $\mathbf{A}_f$ on the right shows the split into two factions with positive ties within each faction and negative ties between factions. It is given by the outer product Eq. (71) with $\mathbf{u}_1 = \mathbf{u}_C$ so that the factions correspond with the identity blocks A and B.

VII. STRUCTURAL BALANCE BEHAVIORAL REGIMES

As the final structure of these dynamical networks is dominated by the initial network's leading eigenvector, we can determine the extent to which networks in our parameter space will become assortative or homogeneous using our transition formulas derived above. First we treat a special case before exploring more general parameters.

A. Ingroup affinity equals outgroup animosity

We consider the simple case in which the ingroup affinity is set equal to the outgroup animosity, $p_{in}^+ = p_{out}^-$. We also make the simplification $d_{in} = d_{out} = d$. For this case, $\langle A_{in} \rangle = -\langle A_{out} \rangle$ so that $\mu = 0$, i.e., there is no homogeneous signal, and $\nu = -\langle A_{out} \rangle$. Using $\langle A_{out} \rangle = d(1 - 2p_{out}^-)$ in Eq. (63), we

solve for the critical outgroup animosity

$$p_{out}^- = \frac{1}{2}\left(1 + \sqrt{\frac{1}{dN}}\right). \tag{72}$$

Note that we only use the positive sign from Eq. (63), since it is the assortative transition, not the disassortative one, that involves the leading eigenvector.

Figure 7 shows the alignment between the dynamical regimes evolved by the structural balance model and the community structure of the initial network. Figure 7(a) plots the assortativity r defined by Eq. (6) of the final adjacency matrix, $\mathbf{A}_f$ as averaged over 400 initial networks generated by the stochastic block model at each point in the d and p_{out}^- parameter space. The yellow region represents the fully assortative outcome where the system evolves into two factions corresponding to the identities defined by the stochastic block model. In the blue region where $r \approx 0$, the two final factions are well mixed by identity. We see that the boundary between these two regions is in good accord with Eq. (72). As the network density increases, the assortative regime is observed to grow, extending down to smaller outgroup animosity values. Figures 7(b) and 7(c) plot, respectively, the spectrum and first eigenvector of the initial adjacency matrix, $\mathbf{A}_0$ for a constant density value. The transition in the structural balance dynamics mirrors the behavior of the first eigenvector which undergoes an assortative detectability transition at $p_{out}^- = 0.525$. These plots also confirm that the

leading eigenvector is never homogeneous and so one does not expect to observe the single faction outcome in this case (being extremely improbable).

B. General parameter conditions

We now analyze the final states of networks generated using more general parameter conditions. We plot the behavior of the signal transition curves in the two-dimensional parameter space defined by p_{in}^+ and p_{out}^- for fixed values of d_{in}/d_{out}.

As the homogeneous signal was irrelevant in the previous case, we needed only plot the assortativity r. However, the prosocial transition will occur in general and so we must measure the extent to which nodes can be found in one large group. We define the homogeneity h as the fraction of all nodes that can be assigned to a single group by virtue of having a common sign in the leading eigenvector of the adjacency matrix. When all nodes have a common sign, they are positively connected to all other nodes so that $h = 1$, while when nodes are divided into two equal factions, the homogeneity assumes its minimum value, $h = 1/2$.

The top plots of Fig. 8 show the assortativity and homogeneity of the final adjacency matrix evolved by the structural balance dynamics in the parameter space defined by the ingroup affinity and outgroup animosity. They can then be linearly combined to effect their joint visualization as shown in the bottom plot. The assortative transition boundary predicted by Eq. (66) separates the assortative from nonassortative two-faction states while the prosocial theoretical boundary of Eq. (68) separates homogeneous single-faction states from the nonassortative two-faction states. These regimes relate to the initial network spectrum as follows: the blue region is where the homogeneous eigenvalue is both the largest eigenvalue and outside the main band; the cyan region is where the leading eigenvalue is part of the main band; and the yellow is where the contrast eigenvalue is largest and outside the main band.

The horizontal yellow-blue interface observed in Fig. 8(c) for larger p_{in}^+ values corresponds to the signal crossing transition in which the homogeneous and contrast eigenvectors exchange places, which occurs outside the noise band (see Fig. 4). Equating the contrast and homogeneous eigenvalue expressions, Eqs. (57) and (58), we find that the transition occurs when $\nu = \mu$, which implies that $\langle A_{out} \rangle = 0$ or equivalently $p_{out}^- = 0.5$.

Figure 9 shows how the density ratio d_{in}/d_{out} and overall network density $d = (d_{in} + d_{out})/2$ affect the assortative and prosocial transition curves. As d_{in}/d_{out} increases, the transition curves become steeper, implying that denser regions of the connectivity network have more influence on the final structure than sparse regions. Figure 10 shows how network size affects the location and shape of the assortative and homogeneous transitions. As N increases, the transitions become sharper and more closely aligned with the theoretical prediction for the critical value of p_{out}^-.

VIII. DISCUSSION

In this section, we first make some observations concerning our results on community and, more broadly, network structure, a subject of relevance to both unsigned and signed
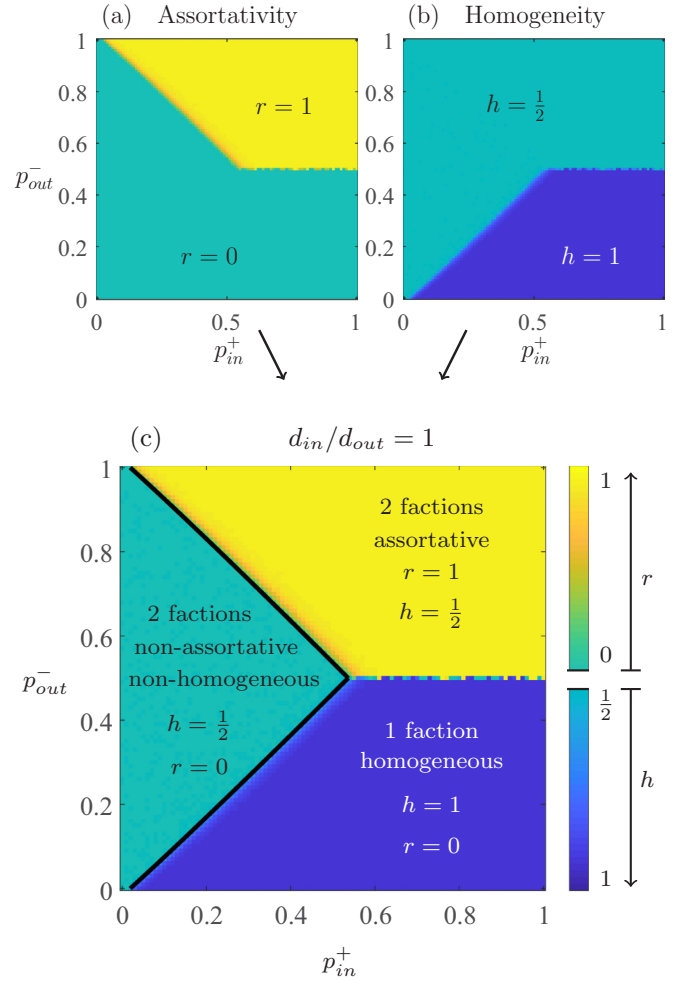

FIG. 8. (a) Assortativity and (b) homogeneity of final network states evolved by structural balance model Eq. (69) as a function of ingroup affinity and outgroup animosity. (c) Assortativity and homogeneity are mapped using the measure $z = r - 2h + 1$ to generate the joint heat map. For convenience, two separate color-bar scales are shown instead of z. The upper black curve indicates the assortative transition boundary, Eq. (66), while the lower black curve indicates the homogeneous transition boundary, Eq. (68). Heatmap values generated by averaging over four simulations for parameters $d_{in} = d_{out} = 0.45$ and $N = 1000$.

networks. We then turn to structural balance dynamics, an intrinsically signed network avenue of research. As it is particularly applicable to social systems, we speculate as to connections between our results and the dynamics of conflicts.

Our results can be applied to unsigned networks by taking the ingroup and outgroup animosities to be zero so that $p_{in}^+ = 1$ and $p_{out}^+ = 1$. In this case, our expressions for the homogeneous and contrast eigenvalues can be reduced to the the sparse-limit forms reported in Refs. [13,14] for the two leading eigenvalues of the adjacency matrix. These can be obtained by neglecting the μ^2 and ν^2 contributions to σ^2 in Eq. (24) and then inserting into Eqs. (58) and (57) for the homogeneous and contrast eigenvalues, respectively. This yields the expressions (outside the noise band),

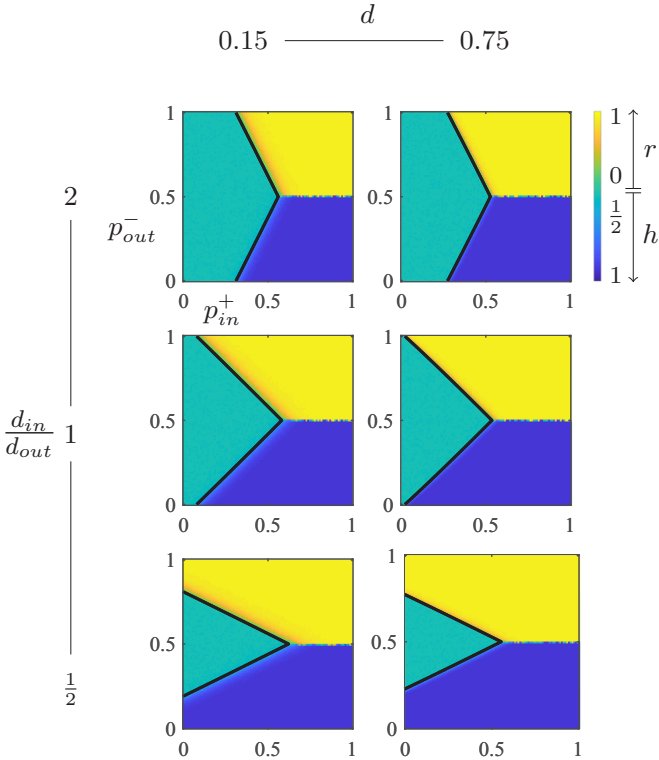$$\lambda_H = \frac{N(d_{in} + d_{out})}{2} + 1, \tag{73}$$

FIG. 9. Final state heat maps for increasing network density $d = (d_{\text{in}} + d_{\text{out}})/2$ and decreasing density ratio $d_{\text{in}}/d_{\text{out}}$. The assortativity and homogeneity are integrated via the z metric (see Fig. 8). Heatmap values generated by averaging over 4 simulations with $N = 1000$.

$$\lambda_C = \frac{N(d_{\text{in}} - d_{\text{out}})}{2} + \frac{d_{\text{in}} + d_{\text{out}}}{d_{\text{in}} - d_{\text{out}}}. \tag{74}$$

These forms, however, do not manifest the μ, ν interchange symmetry, an essential feature of the signed network case.

An important difference between unsigned and signed networks concerns which of the two leading eigenvectors outside the main spectral band may signify community structure. The above equations can be used to show that we can never observe the contrast eigenvalue as being larger than the homogeneous eigenvalue in the unsigned case. The sign of the difference, $\lambda_H - \lambda_C$, depends on the sign of $N - 2/(d_{\text{in}} - d_{\text{out}})$. The second term is equal to $1/\nu$, so that we must have $\nu < 1/N$ for the contrast eigenvalue to exceed the homogeneous eigenvalue. This condition upon ν in conjunction with Eq. (63) for the assortative transition in the sparse limit, which sets the minimum value of ν for λ_C to appear outside the noise band, then necessitates $\mu = (d_{\text{in}} + d_{\text{out}})/2 < 1/N$. However, this regime is below the threshold, $\mu_* = 1/N$, for the prosocial transition in the sparse limit as obtained from Eq. (67). Further, the righthand sides of Eqs. (63) and (67) are the same so that the assortative and prosocial transitions occur at the same critical value, $\nu_* = \mu_*$. For unsigned networks, $\nu \leqslant \mu$ and so if $\mu < \mu_*$, then $\nu < \nu_*$. Therefore, if λ_H is within the noise band then so must λ_C and hence we can never observe $\lambda_C > \lambda_H$.

Consequently, in unsigned networks, assortative community structure is represented by the second eigenvalue of the
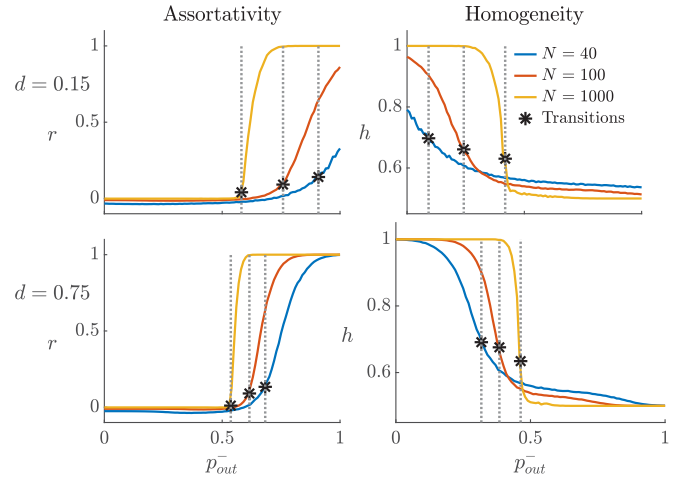

FIG. 10. Assortativity and homogeneity as a function of p_{out}^- for networks of increasing size. Black stars mark the predicted values of p_{out}^- at the transition points as given by Eqs. (66) and (68) for the assortative and prosocial transitions, respectively. Parameter values are $p_{\text{in}}^+ = 0.5$, $d = 0.15$, 0.75, and $N = 40$, 100, 1000 averaged over 2000, 1000, and 10 trials, respectively.

adjacency matrix (when past the detectability threshold) but not the first. In signed networks, the ordering of μ and ν is not restricted and so the first eigenvalue may signify community structure while the second, if above the prosocial transition, signifies the homogeneous structure. Relatedly, while the number of outlying eigenvalues is equal to the number of communities in unsigned networks [8], this need not be the case in signed networks. For instance, the case of equal ingroup affinity and outgroup animosity treated in Sec. VII has only one outlying eigenvalue but two communities.

The sociality transitions, which involve the homogeneous eigenvector, bear upon the question of whether a network exhibits a propensity toward positive versus negative tie formation, a question that is unique to signed networks. In unsigned networks, the prosocial transition is present but its significance corresponds to the emergence of a giant connected component; Eq. (67) yields a transition condition of $N(d_{\text{in}} + d_{\text{out}})/2 = 1$ in accordance with the result for a giant component with two-community structure noted in Sec. II A. Typically, community structure is taken to connote the existence of multiple communities as it is linked to the community detection problem and the assignment of nodes to communities. The sociality transitions, which are not relevant to the community assignment problem, involve network structure more generally rather than community structure per se. However, the sociality transitions do lend themselves to a sense by which a signed network can be viewed as forming a single community: the existence of a global tendency, irrespective of ingroup and outgroup distinctions, toward the formation of positive versus negative ties. The prosocial transition provides a spectral signature for concluding that the network forms a single community in which relationships are generally friendly or cooperative. Conversely, the antisocial transition provides a signature of a single (anti)community marked by hostility, a Hobbesian

state of all against all. The intermediate case, where the homogeneous eigenvalue does not appear outside the noise band, can be taken as marking the absence of a single community with a definitive inclination toward positive or negative tie formation.

Turning to the dynamics, taking networks generated by the stochastic block model and evolving them under the structural balance model entails a dramatic shift in the processes governing network evolution (beyond the change from stochastic to deterministic): a process in which tie formation in a dyad depends only upon its ingroup or outgroup identity is replaced by one where the friendly or hostile relationships among mutual neighbors drives tie value evolution regardless of identity. One rationale for such a shift can be provided by assuming that there is a qualitative change in the nature of the interactions. For instance, hostile relationships characterized by insulting words or gestures may be replaced by physical violence. A second rationale could involve, not a change in the nature of interactions, but a growing awareness that hostile interactions have the potential to become much more prevalent. For example, in a country or region containing two broad identity types, such as ethnicity or ideology, in which individuals or small gangs sporadically clash (either within or across identity lines), a sudden collapse of the central government may lead to a growing sense of looming systemic violence. In either of these rationales, nodes are motivated to seek and maintain allies so that another node's status as the enemy of an enemy or friend of a friend becomes a crucial determinant in relationship formation and evolution.

Our results may inform debates about the interplay of identity and power in conflicts under anarchy consisting of many actors such as insurgencies, civil wars, and international relations. For ethnic conflict, the shift in models discussed above is supported by the observation that the turn from nonviolent to violent conflict represents a qualitative change in dynamics [29]. In the literature on civil wars and ethnic violence, some theories stress mechanisms in which ethnic or religious identity plays an intrinsic role in producing high levels of polarization and violence along identity lines while other theories stress the role of microprocesses of conflict among local actors rather than a pre-existing identity schism [29–31]. As node identity plays no role in the network evolution, structural balance dynamics is consistent with the latter view. However, the sharp transition to the assortative state shows that the dynamics can lock in initial differences in identity-driven community structure even when they are not large, a behavior consistent with the observation that the polarization and violence in ethnic civil wars often appears to be disproportionate to the initial level of ethnic tension. But the existence of the nonassortative regime implies that identity polarization will not arise when the initial structure is sufficiently weak. Thus, conflict takes on an essentially binary nature in that it is either completely polarized by identity or not at all. Additionally, it has been argued that, contrary to some theories, there is no inherent difference in the dynamics between ideological and ethnic civil wars in terms of their potential for polarization and violence [30]. Our results are consistent with such a claim as it is the initial relationships that matter regardless of whether they are due to similar ethnicity or similar ideology.

IX. CONCLUSION

This paper has contributed to two distinct areas of signed network research—community structure and structural balance theory, linking them via the impact of the former upon the latter. We have elucidated the spectrum of unweighted and undirected signed networks generated by a two-community stochastic block model via two independent methods, perturbation analysis and, in the Appendices, a random matrix theory treatment that extends prior work on unsigned networks [13,14]. The expected matrix, $\langle \mathbf{A} \rangle$, in the block model can be decomposed into two signals—a homogeneous eigenvector, $\mathbf{u}_H$, related to the expected tie value, μ, over the network and a contrast eigenvector, $\mathbf{u}_C$, related to the half-difference, ν, between the expected ingroup and outgroup tie values and which encodes the community structure. These signal eigenvectors exhibit transitions at the points where they merge with the main spectral band associated with the noise produced by the zero-mean random matrix $\mathbf{X}$. There are four potential transitions corresponding to the intersections of the two signals with the positive and negative edges of the main band. For the contrast eigenvector, these intersections induce the assortative and disassortative transitions, respectively, and mark changes in community detectability. The homogeneous eigenvector undergoes sociality transitions, prosocial and antisocial, in which emergence from the noise band signifies an overall tendency toward the formation of cooperative or conflictual relationships, respectively, with other nodes.

We derived analytical expressions for the signal eigenvalues in the presence of the noise by performing a perturbation expansion in which the contributions from the noise $\mathbf{X}$ were treated as small corrections to the eigenvalues of $\langle \mathbf{A} \rangle$. Equations (26) and (27) reveal a second-order correction proportional to the average tie variance and are symmetric under the interchange of μ and ν. The same expressions are derived in the Appendices using random matrix theory along with the formula for the main band edges that is a straightforward modification of Wigner's semicircle law. The transition conditions, Eqs. (60) and (61), were determined by equating the signal eigenvalues to the band edge eigenvalues.

We investigated structural balance dynamics in the presence of initial community structure generated by the two-identity stochastic block model. These dynamics completely connect all nodes and allow for three broad regimes of final states: an assortative regime in which two hostile factions emerge that completely align with the two identity blocks; a nonassortative regime in which the two final factions are randomly composed with respect to identity; and a homogeneous regime consisting of a single harmonious faction with only positive ties. Since the dominant eigenvector of the initial network drives its structural balance dynamics and determines its final state, our spectral analysis allows us to chart the parameter conditions under which each of these states will emerge. The dynamical ascendance of the leading eigenvector implies that the regime boundaries occur where any two of the homogeneous signal, the contrast signal, and the noise band edge exchange places as the first eigenvalue. The assortative transition marks the boundary between the nonassortative and assortative regimes and the prosocial transition divides the nonassortative and homogeneous regimes. The boundary

between the homogeneous and assortative regimes represents a reversal in the ranking of the homogeneous and contrast eigenvalues rather than a transition involving the noise band. The theoretically predicted boundaries were found to agree with the simulation results obtained by solving the structural balance model over many random initial networks.

Finally, we note a few potential directions for future research. As with the unsigned case, spectral analysis of signed networks with community structure could be extended to systems with multiple communities, directed ties, and more realistic network statistics such as nonuniform degree distributions. The structural balance model we used is very simple and, problematically, leads to tie strengths which blow up in finite time. Accordingly, the extent to which the dynamical transitions we have identified persist for more realistic implementations of structural balance dynamics should be explored. More empirical work is also needed to understand the conditions under which real networks can be reasonably modeled by structural balance dynamics or variants thereof.

Code reproducing select results from this paper is available online [32].

APPENDIX A: SPECTRUM OF RANDOM MATRIX X

The signal eigenvalues of $\mathbf{A}$ are only visible if they are distinguishable from the noise in the system. In this section we find $\mathcal{O}(||\mathbf{X}||_2)$ as well as γ the spectral edge of the noise matrix $\mathbf{X}$ using random matrix theory.

We start by characterizing the distribution of eigenvalues. The empirical spectral distribution (e.s.d.) of a random matrix $\mathbf{X}$ is defined by

$$\rho(z) = \frac{1}{N} \sum_{i=1}^{n} \delta(z - \omega_i), \tag{A1}$$

where ω_i are the eigenvalues of $\mathbf{X}$.

Wigner's semicircle distribution Eq. (A2) defines the eigenvalue density function for a symmetric random matrix size N with i.i.d. entries having variance m^2,

$$\phi(z) = \frac{1}{2\pi N m^2} \sqrt{4Nm^2 - z^2}. \tag{A2}$$

According to the Wigner limit theorem, the spectra of certain symmetric random matrices converge in distribution to Wigner's semicircle distribution [33,34], so that

$$\lim_{N \to \infty} \int_{-\infty}^{c} \rho(x)dx = \int_{-\infty}^{c} \phi(x)dx, \tag{A3}$$

where c is any real number. Although the semicircle law originally applies only to random symmetric matrices with equal variances for all entries

[33,35,36], further inquiry has determined that random symmetric block Toeplitz matrices [37,38] also weakly converge to the semicircle law under certain conditions.

$\mathbf{X}$ is a random symmetric block Toeplitz matrix with distribution variances σ_{in}^2 and σ_{out}^2 in the on and off diagonal blocks, respectively, and therefore has a Wigner semicircle distribution of eigenvalues for some variance parameter m^2 which we have yet to determine. The edges of the semicircle enclose the band of eigenvalues in the interval $(-2m\sqrt{N}, 2m\sqrt{N})$. As $N \to \infty$, $\lambda_1(\mathbf{X}) = 2m\sqrt{N}$.

We find the variance parameter by bounding $||\mathbf{X}||_2$ with the Frobenius norm [39],

$$\frac{1}{\sqrt{N}} ||\mathbf{X}||_F \leqslant ||\mathbf{X}||_2 \leqslant ||\mathbf{X}||_F, \tag{A4}$$

where

$$||\mathbf{X}||_F = \sqrt{\sum_{j=1}^{\frac{N^2}{2}} |X_{\text{in}}|^2 + \sum_{j=1}^{\frac{N^2}{2}} |X_{\text{out}}|^2}. \tag{A5}$$

The ingroup and outgroup sums are distributed as

$$\sum_{j=1}^{\frac{N^2}{2}} |X_{\text{in}}|^2 \sim \mathcal{N}\left(\frac{N^2}{2}\sigma_{\text{in}}^2, \frac{N^2}{2}\text{Var}(|X_{\text{in}}|^2)\right), \tag{A6}$$

$$\sum_{j=1}^{\frac{N^2}{2}} |X_{\text{out}}|^2 \sim \mathcal{N}\left(\frac{N^2}{2}\sigma_{\text{out}}^2, \frac{N^2}{2}\text{Var}(|X_{\text{out}}|^2)\right). \tag{A7}$$

The random variable Z_2 denoting the inside of the square root in Eq. (A5) is therefore distributed as

$$Z_2 = \sum_{j=1}^{\frac{N^2}{2}} |X_{\text{in}}|^2 + \sum_{j=1}^{\frac{N^2}{2}} |X_{\text{out}}|^2 \sim \mathcal{N}(N^2\sigma^2, N^2\xi^2), \tag{A8}$$

$$\text{where } \xi^2 = \frac{\text{Var}(|X_{\text{in}}|^2) + \text{Var}(|X_{\text{out}}|^2)}{2}. \tag{A9}$$

The expected value of Z_2 scales with $\sigma^2 N^2$ while the standard deviation only scales with ξN meaning $Z_2 = \mathcal{O}(\sigma^2 N^2)$. Therefore,

$$||\mathbf{X}||_F = \mathcal{O}(\sigma N). \tag{A10}$$

Equation (A4) then implies that

$$\sigma\sqrt{N} \leqslant ||\mathbf{X}||_2 \leqslant \sigma N. \tag{A11}$$

Therefore, $||\mathbf{X}||_2$ must scale with σ as

$$\mathcal{O}(||\mathbf{X}||_2) \sim \sigma. \tag{A12}$$

This result combined with the Wigner's semicircle distribution scaling implies that

$$\mathcal{O}(||\mathbf{X}||_2) = \sigma\sqrt{N}. \tag{A13}$$

We continue the argument to find the leading eigenvalue of $\mathbf{X}$. We have found that our variance parameter must scale with σ, $m = \mathcal{O}(\sigma)$,

$$m = a\sigma = a\sqrt{\frac{\sigma_{\text{in}}^2 + \sigma_{\text{out}}^2}{2}}, \tag{A14}$$

for some constant a.

We set $\sigma_{\text{in}}^2 = \sigma_{\text{out}}^2$ which implies that we have a traditional Wigner matrix with i.i.d. entries of variance σ_{in}^2. This means $m = \sigma_{\text{in}} = \sigma$ and therefore $a = 1$. Now let $\sigma_{\text{in}}^2 \neq \sigma_{\text{out}}^2$ while keeping σ^2 constant. m scales only with σ and therefore must be still equal to σ. This means that for all variance values, $m = \sigma$.

This implies that the eigenvalue density function of $\mathbf{X}$ with a diagonal block variance of σ_{in}^2 and an off diagonal block variance of σ_{out}^2 is equivalent to the eigenvalue density function of a random matrix with uniform variance σ^2. This substitution is made in previous derivations of the spectral band of unsigned stochastic block model matrices [13].

Thus, we have determined that the spectra of $\mathbf{X}$ has a Wigner's semicircle distribution with variance parameter σ. Therefore, the edge of the spectral band γ of $\mathbf{X}$ is

$$\gamma = 2\sigma\sqrt{N}. \tag{A15}$$

APPENDIX B: SPECTRA OF A DERIVED FROM RANDOM MATRIX THEORY AND COMPLEX ANALYSIS

We will now use an alternative method to derive the spectra of $\mathbf{A}$ using random matrix theory and complex analysis. In this argument, we use the eigenvalues of the noise matrix $\mathbf{X}$, whose spectra we have defined in Appendix A, to find the eigenvalues of $\mathbf{X} + \nu N \mathbf{u}_C \mathbf{u}_C^T$. We then take these intermediate eigenvalues and use them to find the eigenvalues of $\mathbf{A} = \mathbf{X} + \nu N \mathbf{u}_C \mathbf{u}_C^T + \mu N \mathbf{u}_H \mathbf{u}_H^T$.

The expected adjacency matrix $\langle \mathbf{A} \rangle$ is given by Eq. (13). We consider the spectrum obtained by adding just the contribution of the contrast eigenvector, $\nu N \mathbf{u}_C \mathbf{u}_C^T$, to the noise matrix which yields the eigenvalue equation

$$\left(\mathbf{X} + \nu N \mathbf{u}_C \mathbf{u}_C^T\right)\mathbf{v} = z\mathbf{v}. \tag{B1}$$

We wish to solve for the eigenvalues z, and so use the methods of Ref. [14] to convert Eq. (B1) into a trace representation Eq. (B6). We begin by rearranging the terms in Eq. (B1) to eliminate the eigenvector $\mathbf{v}$,

$$\mathbf{u}_C^T (z - \mathbf{X})^{-1} \mathbf{u}_C = \frac{1}{\nu N}. \tag{B2}$$

The left-hand side of Eq. (B2) can be written as a sum by performing an eigenvector decomposition on $\mathbf{X}$,

$$\mathbf{u}_C^T \mathbf{S} (z\mathbf{I} - \mathbf{\Lambda})^{-1} \mathbf{S}^T \mathbf{u}_C = \frac{1}{\nu N}, \tag{B3}$$

$$\sum_{i=1}^N \frac{\left(\mathbf{x}_i^T \mathbf{u}_C\right)^2}{z - \omega_i} = \frac{1}{\nu N}, \tag{B4}$$

where $\mathbf{S}\mathbf{\Lambda}\mathbf{S}^T$ is the eigenvector decomposition of $\mathbf{X}$ and $\mathbf{x}_i$ are the eigenvectors of $\mathbf{X}$ (as well as the columns of $\mathbf{S}$). We find that $N(\mathbf{x}_i^T \mathbf{u}_C)^2 \sim \mathcal{X}_1^2$ (the chi-square distribution) and therefore $\mathbb{E}[(\mathbf{x}_i^T \mathbf{u}_C)^2] = \frac{1}{N}$ and $\text{Var}[(\mathbf{x}_i^T \mathbf{u}_C)^2] = \frac{2}{N^2}$. This allows us to make the approximation $(\mathbf{x}_i^T \mathbf{u}_C)^2 = 1/N$ in Eq. (B4), giving us

$$\sum_{i=1}^N \frac{1/N}{z - \omega_i} = \frac{1}{\nu N}, \tag{B5}$$

$$\frac{1}{N}\text{Tr}(z - \mathbf{X})^{-1} = \frac{1}{\nu N}. \tag{B6}$$

We define $f(z)$ as follows:

$$f(z) = \text{Tr}(z - \mathbf{X})^{-1}. \tag{B7}$$

The values of z that satisfy $f(z) = \frac{1}{\nu}$ are the eigenvalues of the matrix $\mathbf{X} + \nu N \mathbf{u}_C \mathbf{u}_C^T$. $f(z)$ has simple poles where $z = \omega_i$, $f(z) \to -\infty$ as $z \nearrow \omega_i$ and $f(z) \to \infty$ as $z \searrow \omega_i$. $f(z)$ is a continuous function within the interval $z \in [\omega_i, \omega_{i-1}]$, therefore for each interval $f(z) = \frac{1}{\nu}$ for some value $z \in [\omega_i, \omega_{i-1}]$. This means the eigenvalues z_i and ω_i are interlaced with the leading eigenvalue $z_1 > \omega_1$. The largest solution to $f(z) = \frac{1}{\nu}$ Eq. (B7) is the leading eigenvalue of $\mathbf{X} + \mu N \mathbf{u}_C \mathbf{u}_C^T$.

We now can repeat this process to find a formula for both leading eigenvalues by adding the homogeneous signal in addition to the contrast signal back into the noise matrix and solving for the resulting eigenvalues λ [14].

The new eigenvalue equation becomes

$$\left(\mathbf{X} + \nu N \mathbf{u}_C \mathbf{u}_C^T + \mu N \mathbf{u}_H \mathbf{u}_H^T\right)\mathbf{v} = \lambda\mathbf{v}. \tag{B8}$$

We solve this equation for the eigenvalues λ_i using the same method used to solve Eq. (B1) and which is detailed in Ref. [14]. The resulting equation $g(\lambda)$ has a similar form to $f(z)$ but with an additional term,

$$\frac{1/N}{\lambda - z_1} + \sum_{i=2}^N \frac{1/N}{\lambda - z_i} = \frac{1}{\mu N}, \tag{B9}$$

$$g(\lambda) = \frac{1}{\lambda - z_1} + \sum_{i=2}^N \frac{1}{\lambda - z_i}. \tag{B10}$$

The values of λ that satisfy $g(\lambda) = \frac{1}{\mu}$ are the eigenvalues of the matrix $\mathbf{X} + \nu N \mathbf{u}_C \mathbf{u}_C^T + \mu N \mathbf{u}_H \mathbf{u}_H^T$. Without loss of generality, assume $\nu > \mu$, meaning we have added the largest eigenvalue mode to the noise matrix followed by the second largest eigenvalue mode. When $|\lambda - z_1| \gg \frac{1}{N}$, $\frac{1}{\lambda - z_1} = \mathcal{O}(1)$ and $\sum_{i=2}^N \frac{1}{\lambda - z_i} = \mathcal{O}(N)$. Because $\frac{1}{\lambda - z_1}$ is the dominant term only when $|\lambda - z_1| = \mathcal{O}(1/N^2)$ and the spectral values z_i that constitute the spectral band of $\mathbf{X} + \nu N \mathbf{u}_C \mathbf{u}_C^T$ are interlaced with the spectrum of $\mathbf{X}$, we may approximate $g(\lambda)$ with our previous function $f(\lambda) = \text{Tr}(\lambda - \mathbf{X})^{-1}$ for all λ values away from z_1. $g(\lambda)$ has a singularity at z_1 meaning there is an additional solution to $g(\lambda) = \frac{1}{\mu}$ when $\lambda \approx z_1$.

Therefore, the signal eigenvalues, λ_H and λ_C are the largest magnitude solutions to $f(\lambda) = \frac{1}{\nu}$ and $f(\lambda) = \frac{1}{\mu}$. We can find an analytical form for $f(\lambda)$ by taking advantage of the Stieltjes transform representation of $Tr(\mathbf{X} - \lambda)^{-1}$.

The Stieltjes transform $S_\rho(\lambda)$ of density $\rho(t)$ is a function of the complex variable λ and is defined outside the real interval I,

$$S_\rho(\lambda) = \int_I \frac{\rho(t)}{\lambda - t} dt, \quad \lambda \in \mathbb{C} \backslash I. \tag{B11}$$

The normalized trace of $(\mathbf{X} - \lambda)^{-1}$ is equivalent to the Stieltjes transform of the spectral density of $\mathbf{X}$ [34],

$$\frac{1}{N}\text{Tr}(\mathbf{X} - \lambda)^{-1} = S_\rho(\lambda) = \int_I \frac{\rho(x)}{x - \lambda} dx, \tag{B12}$$

where the e.s.d. has been previously defined, Eq. (A1). We may substitute the eigenvalue density function $\phi(x)$ for the

e.s.d. integrand in the Stieltjes transform since these functions converge in distribution [34],

$$S_\rho(\lambda) = \int_I \frac{\rho(x)}{x - \lambda} dx = \int_I \frac{\phi(x)}{x - \lambda} dx. \quad \text{(B13)}$$

Substituting Eq. (A2) for $\phi(x)$ yields

$$S_\phi(\lambda) = \int_I \frac{\phi(x)}{x - \lambda} dx = \frac{1}{2\pi N \sigma^2} \int_{-2\sqrt{N}\sigma}^{2\sqrt{N}\sigma} \frac{\sqrt{4N\sigma^2 - x^2}}{x - \lambda} dx. \quad \text{(B14)}$$

We solve this integral using multiple changes of variables. Our argument is adapted from previous work [34]. Letting $x = 2\sqrt{N}\sigma \cos(y)$, the above becomes

$$S_\phi(\lambda) = \frac{1}{\pi} \int_0^{2\pi} \frac{\sin^2(y)}{2\sqrt{N}\sigma \cos(y) - \lambda} dy \quad \text{(B15)}$$

$$= \frac{1}{\pi} \int_0^{2\pi} \frac{\left(\frac{e^{iy} - e^{-iy}}{2i}\right)^2}{2\sqrt{N}\sigma\left(\frac{e^{iy} + e^{-iy}}{2}\right) - \lambda} dy. \quad \text{(B16)}$$

The second change of variables is $\zeta = e^{iy}$, which gives

$$S_\phi(\lambda) = \frac{i}{4\pi\sigma\sqrt{N}} \oint_{|\zeta|=1} \frac{(\zeta^2 - 1)^2}{\zeta^2(\zeta^2 - \frac{\lambda}{\sigma\sqrt{N}}\zeta + 1)} d\zeta, \quad \text{(B17)}$$

Let $h(\zeta) = \frac{(\zeta^2 - 1)^2}{\zeta^2(\zeta^2 - \frac{\lambda}{\sigma\sqrt{N}}\zeta + 1)}.$ \quad (B18)

The function $h(\zeta)$ has three poles,

$$\zeta_0 = 0, \quad \text{(B19)}$$

$$\zeta_1 = \frac{\lambda + \sqrt{\lambda^2 - 4N\sigma^2}}{2\sigma\sqrt{N}}, \quad \text{(B20)}$$

$$\zeta_2 = \frac{\lambda - \sqrt{\lambda^2 - 4N\sigma^2}}{2\sigma\sqrt{N}}. \quad \text{(B21)}$$

We must determine which poles are inside the radius $|\zeta| = 1$ and then use the residue theorem to compute the integral. ζ_0 is a pole of order 2 and is inside the contour. Note that $\zeta_1\zeta_2 = 1$, and therefore if $|\zeta_2| \neq |\zeta_1|$ then only one of these poles can be inside the contour. We find, with the argument to follow, that ζ_2 is the pole inside the contour for λ values for which $\text{Im}(\lambda) > 0$ and ζ_1 is the pole inside the contour for λ values for which $\text{Im}(\lambda) < 0$.

We find that ζ_2 is the enclosed pole for $\text{Im}(\lambda) > 0$ by first supposing that $\text{Re}(\lambda) > 0$. It follows that $\text{Re}(\sqrt{\lambda^2 - 4N\sigma^2}) > 0$ and $\text{Im}(\sqrt{\lambda^2 - 4N\sigma^2}) > 0$, and therefore $|\lambda - \sqrt{\lambda^2 - 4N\sigma^2}| < |\lambda + \sqrt{\lambda^2 - 4N\sigma^2}|$, which reveals that $|\zeta_2| < |\zeta_1|$. Let us now suppose that $\text{Re}(\lambda) < 0$. It follows that $\text{Re}(\sqrt{\lambda^2 - 4N\sigma^2}) < 0$ and $\text{Im}(\sqrt{\lambda^2 - 4N\sigma^2}) > 0$, and therefore, by the same argument as above, once again $|\zeta_2| < |\zeta_1|$, meaning that ζ_2 is the enclosed pole.

We use a repetitive argument to show that ζ_1 is the enclosed pole for $\text{Im}(\lambda) < 0$. Supposing that $\text{Re}(\lambda) < 0$, we find that $\text{Re}(\sqrt{\lambda^2 - 4N\sigma^2}) > 0$ and $\text{Im}(\sqrt{\lambda^2 - 4N\sigma^2}) > 0$ which informs us that $|\zeta_1| < |\zeta_2|$. Now supposing that $\text{Re}(\lambda) > 0$, we find that $\text{Re}(\sqrt{\lambda^2 - 4N\sigma^2}) < 0$ and $\text{Im}(\sqrt{\lambda^2 - 4N\sigma^2}) > 0$ and therefore $|\zeta_1| < |\zeta_2|$, meaning that ζ_1 is the enclosed pole.

We now know that ζ_1 is inside the contour and ζ_2 is outside for $\text{Im}(\lambda) < 0$ while ζ_2 is inside the contour and ζ_2

is outside for $\text{Im}(\lambda) > 0$. We use the Residue theorem and the three poles Eqs. (B19)–(B21) to finish solving Eq. (B17) by integrating

$$\oint_{|\zeta|=1} h(\zeta) d\zeta = 2\pi i [\text{Res}(h, \zeta_0) + \text{Res}(h, \zeta_*)], \quad \text{(B22)}$$

$$\text{where } \zeta_* = \begin{cases} \zeta_1, & \text{for } \text{Im}(\lambda) < 0, \\ \zeta_2. & \text{for } \text{Im}(\lambda) > 0, \end{cases} \quad \text{(B23)}$$

Evaluating the residues yields

$$\oint_{|\zeta|=1} h(\zeta) d\zeta = 2\pi i \left[\frac{\lambda}{\sigma\sqrt{N}} \pm \frac{\sqrt{\lambda^2 - 4N\sigma^2}}{\sigma\sqrt{N}} \right]. \quad \text{(B24)}$$

The $\pm$ in the above expression results from the two different residues for $\text{Im}(\lambda) > 0$ and $\text{Im}(\lambda) < 0$. We now multiply our result by the constant term to finish solving Eq. (B17),

$$S_\phi(\lambda) = \frac{-\lambda \pm \sqrt{\lambda^2 - 4N\sigma^2}}{2N\sigma^2}. \quad \text{(B25)}$$
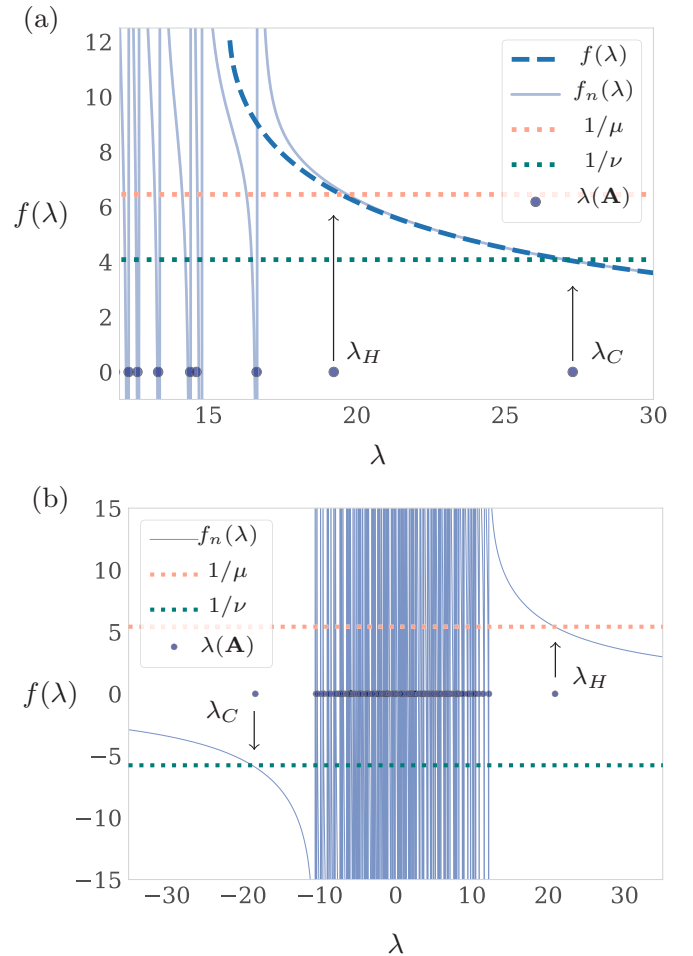

FIG. 11. Intersections of the function $f(\lambda)$ and $1/\mu$ and $1/\nu$ yield eigenvalue locations. (a) $f_n(\lambda)$ is the numerical solution to $f(\lambda)$ and intersects with $1/\mu$ and $1/\nu$ at λ_H and λ_C. $f(\lambda)$ is only defined away from the spectral band; the analytical solution diverges from the numerical approximation upon approaching the spectral edge. (b) The case where λ_H and λ_C are on opposite sides of the spectral band as the signal eigenvalues can be either positive or negative in signed networks.

Now we have a closed form solution for the Stieltjes transform of the spectral density of $\mathbf{X}$ which gives us an analytical formula for $f(\lambda)$ in the region when λ is outside of the spectral band, $|\lambda| > 2\sigma\sqrt{N} = \gamma$,

$$f(\lambda) = \mathrm{Tr}(\lambda - \mathbf{X})^{-1} = \frac{\lambda \pm \sqrt{\lambda^2 - 4N\sigma^2}}{2\sigma^2}, \quad |\lambda| > \gamma. \tag{B26}$$

We now solve for the eigenvalue λ of $\mathbf{A}$ that corresponds to the contrast signal,

$$f(\lambda) = \frac{1}{\nu}, \tag{B27}$$

$$\frac{\lambda \pm \sqrt{\lambda^2 - 4N\sigma^2}}{2\sigma^2} = \frac{1}{\nu} \tag{B28}$$

$$\Rightarrow \lambda = \nu N + \frac{\sigma^2}{\nu}, \quad |\nu| \geqslant \frac{\sigma}{\sqrt{N}}. \tag{B29}$$

If $|\nu| \geqslant \frac{\sigma}{\sqrt{N}}$, then $|\lambda| \geqslant |\gamma|$, and the contrast eigenvalue λ is outside the spectral band, otherwise the leading eigenvalue is included in the spectral band. The analogous formula for the eigenvalue corresponding to the homogeneous signal is found by replacing ν by μ in Eq. (B29). Figure 11 shows how the intersections between $f(\lambda)$ and $1/\mu$ and $1/\nu$ generate the eigenvalue locations. We have therefore found the same solutions for the signal eigenvalues of $\mathbf{A}$ using random matrix theory and complex analysis as done via perturbation theory in the main text, Eqs. (57) and (58).

[1] A. Kirkley, G. T. Cantwell, and M. E. J. Newman, Phys. Rev. E **99**, 012320 (2019).

[2] J. Bruggeman, V. A. Traag, and J. Uitermark, Amer. Sociol. Rev. **77**, 1050 (2012).

[3] G. Facchetti, G. Iacono, and C. Altafini, Proc. Nat. Acad. Sci. USA **108**, 20953 (2011).

[4] Z. Maoz, L. G. Terris, R. D. Kuperman, and I. Talmud, J. Politics **69**, 100 (2007).

[5] P. Doreian and A. Mrvar, J. Soc. Struct. **16**, 1 (2015).

[6] S. Fortunato and D. Hric, Phys. Rep. **659**, 1 (2016).

[7] M. E. J. Newman, Phys. Rev. E **74**, 036104 (2006).

[8] S. Chauhan, M. Girvan, and E. Ott, Phys. Rev. E **80**, 056114 (2009).

[9] V. A. Traag and J. Bruggeman, Phys. Rev. E **80**, 036115 (2009).

[10] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. De Luca, and S. Albayrak, in *Proceedings of the SIAM International Conference on Data Mining* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2010), pp. 559–570.

[11] P. Esmailian and M. Jalili, Sci. Rep. **5**, 14339 (2015).

[12] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, Phys. Rev. Lett. **107**, 065701 (2011).

[13] R. R. Nadakuditi and M. E. J. Newman, Phys. Rev. Lett. **108**, 188701 (2012).

[14] X. Zhang, R. R. Nadakuditi, and M. E. J. Newman, Phys. Rev. E **89**, 042816 (2014).

[15] A. Ghasemian, P. Zhang, A. Clauset, C. Moore, and L. Peel, Phys. Rev. X **6**, 031005 (2016).

[16] M. Wilinski, P. Mazzarisi, D. Tantari, and F. Lillo, Phys. Rev. E **99**, 042310 (2019).

[17] E. Abbe, J. Mach. Learn. Res. **18**, 6446 (2017).

[18] K. Kulakowski, P. Gawroński, and P. Gronek, Int. J. Mod. Phys. C **16**, 707 (2005).

[19] S. A. Marvel, J. Kleinberg, R. D. Kleinberg, and S. H. Strogatz, Proc. Natl. Acad. Sci. USA **108**, 1771 (2011).

[20] M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. USA **99**, 7821 (2002).

[21] M. Newman, *Networks*, 2nd ed. (Oxford University Press, Oxford/New York, 2018).

[22] B. Karrer and M. E. J. Newman, Phys. Rev. E **83**, 016107 (2011).

[23] P. Erdös and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci **5**, 17 (1960).

[24] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).

[25] C. Sarkar and S. Jalan, Chaos **28**, 102101 (2018).

[26] D. Cartwright and F. Harary, Psychol. Rev. **63**, 277 (1956).

[27] V. Traag, P. Doreian, and A. Mrvar, Partitioning signed networks, in *Advances in Network Clustering and Blockmodeling*, edited by P. Doreian, V. Batagelj, and A. Ferligoj (Wiley Online Library, 2019), pp. 225–249.

[28] V. A. Traag, P. V. Dooren, and P. D. Leenheer, PLoS ONE **8**, e60063 (2013).

[29] R. Brubaker and D. D. Laitin, Annu. Rev. Sociol. **24**, 423 (1998).

[30] S. N. Kalyvas and M. A. Kocher, Politics Soc. **35**, 183 (2007).

[31] L.-E. Cederman and M. Vogt, J. Conflict Resolut. **61**, 1992 (2017).

[32] https://github.com/mmtree/Community_signed_networks.

[33] E. P. Wigner, Ann. Math. **67**, 325 (1958).

[34] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Springer Series in Statistics (Springer, New York, NY, 2010).

[35] Z. D. Bai and Y. Q. Yin, Ann. Prob. **16**, 863 (1988).

[36] L. Erdös, Russian Math. Surveys **66**, 507 (2011).

[37] Y.-T. Li, D.-Z. Liu, and Z.-D. Wang, J. Theor. Prob. **24**, 1063 (2011).

[38] R. Basu, A. Bose, S. Ganguly, and R. S. Hazra, Stat. Prob. Lett. **82**, 1430 (2012).

[39] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins Studies in the Mathematical Sciences (Johns Hopkins University Press, Baltimore, MD, 1996).