Continuous-time model of structural balance
Seth A. Marvel, Jon Kleinberg, Robert D. Kleinberg, and Steven H. Strogatz

*Center for Applied Mathematics, and Department of Computer Science, Cornell University, Ithaca, NY 14853

Edited by Ronald L. Graham, University of California, San Diego, La Jolla, CA, and approved November 30, 2010 (received for review September 3, 2010)

It is not uncommon for certain social networks to divide into two opposing camps in response to stress. This happens, for example, in networks of political parties during winner-takes-all elections, in networks of companies competing to establish technical standards, and in networks of nations faced with mounting threats of war. A simple model for these two-sided separations is the dynamical system \( dX/dt = X^2 \), where \( X \) is a matrix of the friendliness or unfriendliness between pairs of nodes in the network. Previous simulations suggested that only two types of behavior were possible for this system: Either all relationships become friendly or two hostile factions emerge. Here we prove that for generic initial conditions, these are indeed the only possible outcomes. Our analysis yields a closed-form expression for faction membership as a function of the initial conditions and implies that the initial amount of friendliness in large social networks (started from random initial conditions) determines whether they will end up in intractable conflict or global harmony.

The mathematical model that we want to study is best understood as an outgrowth of a theory from social psychology known as structural balance (1). So let us begin with a brief explanation of what this theory says.

Consider three individuals: Anna, Bill, and Carl, and suppose that Bill and Carl are friends with Anna, but are unfriendly with each other. If the sentiment in the relationships is strong enough, Bill may try to strengthen his friendship with Anna by encouraging her to turn against Carl, and Carl might likewise try to convince Anna to terminate her friendship with Bill. Anna, for her own part, may try to bring Bill and Carl together so they can reconcile and become friends. In abstract terms, relationship triangles containing exactly two friendships are prone to transition to triangles with either one or three friendships.

Alternately, suppose that Anna, Bill, and Carl all view each other as rivals. In many such situations, there are incentives for the two people in the weakest rivalry to cooperate and form a working friendship or alliance against the third. In these cases, a single friendship may be prone to appear in a relationship triangle that initially has none.

These two thought experiments suggest a notion of stability, or balance, that can be traced back to the work of Heider (2). Heider’s theory was expanded into a graph-theoretic framework by Cartwright and Harary (3), who considered graphs on \( n \) nodes (representing people, countries, or corporations) with edges signed either positive (+) to denote friendship or negative (−) to denote rivalry. If a social network feels the proper social stresses (those felt by Anna, Bill, and Carl in the examples above), then Cartwright and Harary’s theory predicts that in steady state the triangles in the graph should contain an odd number of positive edges—in other words, three positive edges or one positive edge and two negative edges. We refer to such triangles as balanced, and triangles with an even number of positive edges as unbalanced. Finally, we call a graph complete if it contains edges between all pairs of nodes, and we say that a complete graph with signs on its edges is balanced if all its triangles are balanced. (All graphs in our discussion will be complete.)

As it turns out, these local notions of balance theory are closely related to the global structure of two opposing factions. In particular, suppose that the nodes of a complete graph are partitioned into two factions such that all edges inside each faction are positive and all edges between nodes in opposite factions are negative. (One of these factions may be empty, in which case the other faction includes all the nodes in the graph, and consequently all edges of the network are positive.) Note that this network must be balanced, because each triangle either has all three members in the same faction (yielding three positive edges) or has two members in one faction and the third member in the other faction (yielding one positive edge and two negative ones). In fact, a stronger and less obvious statement is true: Any balanced graph can be partitioned into two factions in this way, with one faction possibly empty (3). As a result, when we speak of balanced graphs, we can equivalently speak of networks with this type of two-faction structure.

Model
Structural balance is a static theory—it posits what a “stable” signing of a social network should look like. However, its underlying motivation is dynamic, based on how unbalanced triangles ought to resolve to balanced ones. This situation has led naturally to a search for a full dynamic theory of structural balance. Yet finding systems that reliably guide networks to balance has proved to be a challenge in itself.

A first exploration of this issue was conducted by Antal et al. (4) who considered a family of discrete-time models. In one of the main models of this family, an edge of the graph is examined in each time step, and its sign is flipped if this produces more balanced triangles than unbalanced ones. Although a balanced graph is a stable point for these discrete dynamics, it turns out that many unbalanced graphs called jammed states are as well (4, 5).

Thus, the natural problem became to identify and rigorously analyze a simple system that could progress to balanced graphs from generic initial configurations. An approach to this problem was taken by Kulakowski et al. (6), who proposed a continuous-time model for structural balance. They represented the state of a completely connected social network using a real symmetric \( n \times n \) matrix \( X \) whose entry \( x_{ij} \) represents the strength of the friendliness or unfriendliness between nodes \( i \) and \( j \) (a positive value denotes a friendly relationship and a negative value an unfriendly one). Note that for a given \( X \), there is a signed complete graph with edge signs equal to the signs of the corresponding elements \( x_{ij} \) in \( X \). We will call \( X \) balanced if this associated signed complete graph is balanced.

Kulakowski et al. considered variations on the following basic differential equation, which they proposed as a dynamical system governing the evolution of the relationships over time:

\[
\frac{dX}{dt} = X^2
\]


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.
See Commentary on page 1751.

1To whom correspondence should be addressed. E-mail: kleiber@cs.cornell.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1013213108/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1013213108
Remarkably, simulations showed that for essentially any initial \( X'(0) \), the system reached a balanced pattern of edge signs in finite time. Writing Eq. 1 directly in terms of the entries \( x_{ij} \) gives a sense of why this differential equation should promote balance:

\[
\frac{dx_{ij}}{dt} = \sum_k x_{ik} x_{kj}.
\]

Notice that \( x_{ij} \) is being pushed in a positive or negative direction based on the relationships that \( i \) and \( j \) have with \( k \). If \( x_{ik} \) and \( x_{kj} \) have the same sign, their product guides the value of \( x_{ij} \) in the positive direction, whereas if \( x_{ik} \) and \( x_{kj} \) have opposite signs, their product guides the value of \( x_{ij} \) in the negative direction. In each case, this is the direction required to balance the triangle \( \{i,j,k\} \).

Note also that Eq. 2 applies for the case that \( i = j \). Although this case is harder to interpret, the monotonic increase of \( x_{ii} \) is being pushed in a positive or negative direction, whereas if \( x_{ii} \) is being pushed in a positive or negative direction, the system reaches a balanced pattern of edge signs in finite time with probability converging to 1 as \( T \) signifies transposition.

The differential equation Eq. 1 is a special case of a general family of equations known as matrix Riccati equations (7). The analysis of the full family is complicated and not fully resolved, but we now show that the special case of concern to us, Eq. 1, has an explicit solution with a form that exposes its connections to structural balance. We proceed as follows. First, we observe that by separation of variables, the solution of the single-variable differential equation \( \dot{x} = x^2 \) (overdetermined by the initial condition \( x(0) = \lambda_k \)) is

\[
\dot{\lambda}_k(t) = \frac{\lambda_k}{1 - \lambda_k t}.
\]

Therefore the diagonal matrix \( D(t) = \text{diag}[\dot{\lambda}_1(t), \dot{\lambda}_2(t), \ldots, \dot{\lambda}_n(t)] \) is the solution of Eq. 1 for the initial condition \( X(0) = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n] \).

Moreover \( Y(t) = QD(t)^2 \) is also a solution of Eq. 1 because \( Y = QDQ^T = Q(D^2)^2 = Q(D^2)^2 = Y^2 \). But \( Y(t) \) has the same initial condition as \( X(t) \) in our original problem: \( Y(0) = QD(0)^2 = X(0) \). So by uniqueness, \( Y(t) = QD(t)^2 \) must be the solution we seek.

Our solution \( X(t) \) can also be written in a different way to mimic the solution of the one-dimensional equation \( \dot{x} = x^2 \). By separating variables, we can expand the denominators of the \( \dot{\epsilon}_k(t) \) functions in powers of \( t \) to rewrite \( X(t) \) as \( X(0) + \frac{X(0)^2}{2} + \frac{X(0)^3}{3!} + \ldots \), or more concisely,

\[
X(t) = X(0)[I - X(0)t]^{-1}.
\]

(Note that the matrices \( X(0) \) and \( I - X(0)t \) commute.) This equation is valid when \( t \) is less than the radius of convergence of every \( \lambda_k \), that is, when \( t < 1/\lambda_1 \) (assuming \( \lambda_1 > 0 \)).

Finally we note that the above method of solving Eq. 1 contains a reduction of the number of dynamical variables of the system from \( n(n-1)/2 \) to \( n \). The \( \langle \rangle \) constants of motion generated by this reduction are just the off-diagonal elements of \( Q^T X(t)Q = D(t) \), or \( \sum_{i<j} \sum_{k<l} q_{ik} x_{kj} q_{jl} q_{lk} = 0 \) for all \( 1 \leq i < j \leq n \). Furthermore, the procedure for reducing \( X(t) \) can be easily generalized to any system of the form \( \dot{X} = f(X) \), where \( f \) is a polynomial of \( X \).

Behavior of the Model: Evolution to a Balanced State

Suppose we randomly select the \( x_{ij}(0) \)'s from a continuous distribution on the real line. Then the \( x_{ii}(t)'/s \) found by numerical integration generally sort themselves in finite time into the sign pattern of two feuding factions. To reformulate this observation as a precise statement and explain why the behavior holds so pervasively, we now solve Eq. 1 explicitly.

Solution to the Model.

The initial matrix \( X(0) \) is real and symmetric by assumption, so we can write it as \( QD(0)Q^T \), where \( D(0) \) is the diagonal matrix with the eigenvalues of \( X(0) \), denoted \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \), as diagonal entries ordered from largest to smallest, and \( Q \) is the orthogonal matrix with the corresponding eigenvectors of \( X(0) \), denoted \( a_1, a_2, \ldots, a_n \), as columns. The superscript \( T \) signifies transposition.

The differential equation Eq. 1 is a special case of a general family of equations known as matrix Riccati equations (7). The analysis of the full family is complicated and not fully resolved, but we now show that the special case of concern to us, Eq. 1, has an explicit solution with a form that exposes its connections to structural balance. We proceed as follows. First, we observe that by separation of variables, the solution of the single-variable differential equation \( \dot{x} = x^2 \) (overdetermined by the initial condition \( x(0) = \lambda_k \)) is

\[
\dot{\lambda}_k(t) = \frac{\lambda_k}{1 - \lambda_k t}.
\]

Therefore the diagonal matrix \( D(t) = \text{diag}[\dot{\lambda}_1(t), \dot{\lambda}_2(t), \ldots, \dot{\lambda}_n(t)] \) is the solution of Eq. 1 for the initial condition \( X(0) = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n] \).

Moreover \( Y(t) = QD(t)^2 \) is also a solution of Eq. 1 because \( Y = QDQ^T = Q(D^2)^2 = Q(D^2)^2 = Y^2 \). But \( Y(t) \) has the same initial condition as \( X(t) \) in our original problem: \( Y(0) = QD(0)^2 = X(0) \). So by uniqueness, \( Y(t) = QD(t)^2 \) must be the solution we seek.

Our solution \( X(t) \) can also be written in a different way to mimic the solution of the one-dimensional equation \( \dot{x} = x^2 \). By separating variables, we can expand the denominators of the \( \dot{\epsilon}_k(t) \) functions in powers of \( t \) to rewrite \( X(t) \) as \( X(0) + X(0)^2/2 + X(0)^3/3! + \ldots \), or more concisely,

\[
X(t) = X(0)[I - X(0)t]^{-1}.
\]

(Note that the matrices \( X(0) \) and \( I - X(0)t \) commute.) This equation is valid when \( t \) is less than the radius of convergence of every \( \lambda_k \), that is, when \( t < 1/\lambda_1 \) (assuming \( \lambda_1 > 0 \)).

Finally we note that the above method of solving Eq. 1 contains a reduction of the number of dynamical variables of the system from \( n(n-1)/2 \) to \( n \). The \( \langle \rangle \) constants of motion generated by this reduction are just the off-diagonal elements of \( Q^T X(t)Q = D(t) \), or \( \sum_{i<j} \sum_{k<l} q_{ik} x_{kj} q_{jl} q_{lk} = 0 \) for all \( 1 \leq i < j \leq n \). Furthermore, the procedure for reducing \( X(t) \) can be easily generalized to any system of the form \( \dot{X} = f(X) \), where \( f \) is a polynomial of \( X \).

Behavior of the Solution.

Let us now examine the behavior of our solution \( X(t) \) to see why in the typical case it splits into two factions in finite time. It turns out that this is the guaranteed outcome if the following three conditions hold (and as we see below, they hold with probability converging to 1 as \( n \) goes to infinity):

1. \( \lambda_1 > 0 \),
2. \( \lambda_1 \neq \lambda_2 \) (and hence \( \lambda_1 > \lambda_3 \)), and
3. all components of \( a_n \) are nonzero.

To see why these conditions imply a split into two factions, observe from Eq. 3 that each \( \dot{\epsilon}_k(t) \) diverges to infinity at \( t = 1/\lambda_k \). Because \( x_{ii}(t) = \sum q_{ik} \dot{\epsilon}_k(t) q_{jk} \), all \( x_{ii}'s \) diverge to infinity when
the $c_k$ with the smallest positive $1/\lambda_k$ does. Under the first and second conditions, this $c_k$ is $c_1$, so the blow-up time $t^*$ of Eq. 1 must be $1/\lambda_1$. To show that the nodes are partitioned into two factions as $X(t)$ approaches $t^*$, let $X(t) = X(t)/||X(t)||$ on the half-open interval $[0,t^*)$, where $||X||$ denotes the Frobenius norm of $X$. The matrix $X(t)$ has the sign pattern of $X(t)$, and as $t$ approaches $t^*$ it converges to the rank-one matrix

$$X^* = Q\text{diag}(1,0,0,...,0)Q^T = c_1c_1^T. \quad [5]$$

Now let $\alpha_{tk}$ denote the value of the $k$th coordinate of $\alpha_1$, and let $S = \{k': \alpha_{tk} > 0\}$ and $T = \{k': \alpha_{tk} < 0\}$. Then $S$ and $T$ partition the node indices $1,2,...,n$ by our condition that $\alpha_1$ has no zero components. From Eq. 5, this partition must correspond to two cliques of friends joined by a complete bipartite graph of unfriendly ties.

The Three Conditions. We now return to the three conditions above. We first show that the second and third hold with probability 1. We then show that the first condition holds with probability converging to 1 as $n$ goes to infinity. Finally, we analyze the behavior of the system in the unlikely event that the first condition does not hold. The fact that the conjunction of all three conditions holds with probability converging to 1 as $n$ grows large justifies our earlier claim that the behavior described above holds for almost all choices of initial conditions.

First we show why the second and third conditions hold with probability 1 so long as the (joint) distribution from which $X(0)$ is drawn is absolutely continuous with respect to Lebesgue measure—in other words, assigns probability zero to any set of matrices whose Lebesgue measure is zero. Our arguments below make use of the following two basic facts:

i. The set of zeros of a nontrivial multivariate polynomial has Lebesgue measure zero, and

ii. The existence of a common root of two univariate polynomials $P$ and $Q$ is equivalent to the vanishing of a multivariate polynomial in the coefficients of $P$ and $Q$ (specifically, it is equivalent to the vanishing of the determinant of the Sylvester matrix of $P$ and $Q$, also called the resultant of $P$ and $Q$).

To show that $\lambda_1 \neq \lambda_2$ with probability 1, let $P$ denote the characteristic polynomial of $X(0)$, and let $Q$ denote the derivative of $P$. Then $X(0)$ has a repeated eigenvalue if and only if $P$ has a repeated root, which it does if and only if $P$ and $Q$ have a common root. This condition is equivalent to the vanishing of the resultant of $P$ and $Q$, which is a multivariate polynomial in the entries of $X(0)$. The polynomial cannot be zero everywhere, because there is at least one symmetric matrix that does not have a repeated eigenvalue. So the set of matrices having a repeated eigenvalue has Lebesgue measure zero.

Similarly, to show that all components of $\alpha_1$ are nonzero, let $P$ denote the characteristic polynomial of $X(0)$ and $P$, the characteristic polynomial of the $(n-1) \times (n-1)$ submatrix $X_i(0)$ obtained by deleting the $i$th row and $i$th column of $X(0)$. It is easy to check that if any eigenvector of $X(0)$ has a zero in its $i$th component, then the vector obtained by deleting that component is an eigenvector of $X_i(0)$ with the same eigenvalue. Consequently, $P$ and $P_i$ must have a common root, implying that the resultant of $P$ and $P_i$ vanishes. This resultant is once again a multivariate polynomial in the entries of $X(0)$, and once again it must be nonzero somewhere because there is at least one symmetric matrix whose eigenvectors all have nonzero entries. Hence, the set of matrices having an eigenvector with zero in its $i$th component has Lebesgue measure zero.

Finally, to determine the likelihood of the first condition, we first must say a bit more about the way that $X(0)$ is selected. Suppose that the off-diagonal $x_{ij}(0)$’s are drawn randomly from a common distribution $F$ and the on-diagonal $x_{ii}(0)$’s are drawn randomly from a common distribution $G$. All selections are independent for $i \leq j$. For $i > j$, we let $x_{ij}(0) = x_{ji}(0)$, so that $X(0)$ is symmetric. For this construction of $X(0)$, Arnold (8) has shown that with the remarkably weak additional assumption that $F$ has a finite second moment, Wigner’s semicircle law holds in probability as $n$ grows to infinity. This in turn implies that $\lambda_1 > 0$ in probability in the same limit.

Moreover, suppose we are in the low-probability case that $\lambda_1 \leq 0$. In this case, the analysis above shows that all the functions $\ell_i(t)$ converge to 0 as $t \to \infty$. Thus, $\lim_{t \to \infty} Q(t) = 0$, and we also have $\lim_{t \to \infty} X(t) = 0$.

Although the entries of $X(t)$ converge to zero when $\lambda_1 \leq 0$, one might still want to know if the sign pattern of $X(t)$ is eventually constant (i.e., remains unchanged for all $t$ above some threshold value) and, if so, what determines this sign pattern. It is possible to answer this question, again assuming the second and third conditions. By expanding the function $\ell_i(t) = \lambda_i/(1 - \lambda_i)$ in powers of $u = 1/t$, we obtain the asymptotic series

$$\ell_i(t) = -u - u^2 \lambda_i^{-1} - O(u^3). \quad [6]$$

which implies

$$X(t) = QD(t)Q^T = -tu - t^2X(0)^{-1} - O(u^2). \quad [7]$$

In the limit of small $u$, the leading-order term of the diagonal entries of $X(t)$ is the linear term, which has negative sign. For the off-diagonal entries of $X(t)$, the leading-order term as $u$ tends to zero is the quadratic term, whose sign matches the sign of the corresponding off-diagonal entry of the matrix $-X(0)^{-1}$.

Behavior of the Model: From Factions to Unification

The analysis in the previous section tells us how to find both the blow-up time $t^*$ and final sign configuration of a network if we know its initial state $X(0)$. However, we might also want to know whether we can characterize the behavior of $X(t)$ in the large-$n$ limit in terms of statistical parameters of $X(0)$. This could, for example, help us forecast the behavior of groups of individuals when collecting complete relationship-level data is not feasible. Clearly if the underlying network is a complete graph, it is not realistic to consider $n$ that are too large, but we find fortunately in simulations that the asymptotic behavior we will derive in this section becomes apparent even for moderate values of $n$ (less than 100). As a result, these large-$n$ results are perhaps most usefully viewed as an approximate guide to what happens in mediumsized groups that are large enough to show predictable collective behavior but for which a completely connected set of relationships is still feasible to maintain.

In this section, we show that there is a transition from final states consisting of two factions to final states consisting of all-positive relations as the “mean friendliness” of $X(0)$ (the mean of the distributions used to generate the off-diagonal entries of $X(0)$) is increased from negative to positive values. This is consistent with the numerical simulations shown in Fig. 1; as noted above, the asymptotic behavior we are studying is already clear in these simulations, which are performed with $n = 90$.

Before discussing the details, we describe how $X(0)$ is selected in this section. We start by adopting the procedure of Füredi and Komlós (9): The elements $x_{ij}(0)$ are drawn independently from distributions $F_{ij}$ with zero mass outside of $[-K,K]$. The off-diagonal $F_{ij}$’s have a common expectation $\mu$ and finite variance $\sigma^2$, whereas the on-diagonal $F_{ii}$’s have a common expectation $\nu$ and variance $\nu^2$. In addition, we require that each off-diagonal distribution $F_{ij}$ be symmetric about $\mu$. Random matrix models of this type have attracted considerable recent interest (see, e.g., refs. 10 and 11), but we need only the basic results of Füredi and Komlós (9) for our purposes, and so we use these in the
development below. We consider the three cases of positive, zero, and negative $\mu$.

**Case 1: $\mu > 0$.** The results of Füredi and Komlós (9) show that when $\mu > 0$, the deviation of $o(t)$ from $(1, 1, \ldots, 1)/\sqrt{n}$ vanishes in probability in the large-$n$ limit. Hence the final state of the system consists of one large clique of friends including all but at most a vanishing fraction of the nodes. Moreover, by assuming a bound on $\sigma$ we can strengthen this statement further: If $\sigma < \mu/2$, then the findings of Füredi and Komlós imply that the final state consists of a single clique offriends, with no negative edges. These observations are consistent with the representative numerical trial shown in Fig. 1A. Moreover, Füredi and Komlós show that the asymptotic behavior of $x_i$ grows like $\mu n + O(1)$, and hence the blow-up time scales like $1/(\mu n)$.

We can gain insight into the behavior of the system for small $t$ using an informal Taylor series calculation: If we rescale time in Eq. 1 by inserting a $1/n$ before the summation, compute the Taylor expansion of $x(t)$ term by term, and then take the expectation of each term, we obtain the geometric series $x(t) = \mu + \mu^2 t + \mu^3 t^2 + \ldots$, or

$$x(t) = \frac{\mu}{1 - \mu t}.$$  \[8\]

With significantly more work, it can be proved that every trajectory $x_i(t)$ has this time dependence on $[0, 1/K]$ in the large-$n$ limit with probability 1 (see SI Text), so we may write

$$\lim_{n \to \infty} x_i(t) = x_i(0) - \mu + \frac{\mu}{1 - \mu} t$$  \[9\]

with prob. 1 for all $t$ in $[0, 1/K]$. Observe that this limit has a blow-up time $t^*$ of $1/\mu$. Because our rescaling of time represents a zooming in or magnification of time by a factor of $n$, this $t^*$ corresponds to a blow-up time asymptotic to $1/(\mu n)$ for the rescaled system, consistent with the results of Füredi and Komlós.

**Case 2: $\mu = 0$.** In the event that the network starts from a mean friendliness of zero, numerical experiments indicate that the system ends up with two factions of equal size in the large-$n$ limit (Fig. 1B). We now prove this to be the case. For the remainder of this discussion, we abbreviate $X(0)$ as $A$ and $x_i(0)$ as $a_i$.

Because the off-diagonal entries of $A$ have symmetric distributions by assumption, we have for any off-diagonal $a_{ij}$ and any interval $S_i$ on the real line that $P(a_{ij} \in S_i) = P(-a_{ij} \in S_i)$. Now let $D$ be a diagonal matrix with some sequence of $+1$ and $-1$ along its diagonal (where the $i$th diagonal entry is denoted by $d_i$). Then the random matrices $A$ and $B = DAD$ are identically distributed, as we now show.

To say that $A$ and $B$ are identically distributed means that for every Borel set of matrices $P(A \in S) = P(B \in S)$. To prove this, it suffices to consider the case in which $S$ is a product of intervals $S_i$, because these product sets generate the Borel $\sigma$-algebra. The upper triangular entries of $A$ are independent, so $P(A \in S) = \prod_{i<j} P(a_{ij} \in S_{ij})$. Similarly $P(B \in S) = \prod_{i<j} P(d_id_j \in S_{ij})$. By the symmetry of the off-diagonal distributions, $\prod_{i<j} P(a_{ij} \in S_{ij}) = \prod_{i<j} P(d_id_j \in S_{ij})$, which gives us $P(A \in S) = P(B \in S)$ as desired. (Note that when $i = j$, the factor $d_id_j$ is 1 so the on-diagonal distributions need not be symmetric.)

Now consider the set $S$ of matrices with an $o(n)$ consisting of all positive components. The above demonstration implies that the probability of choosing an $A$ in this set is the same as choosing an $A$ such that $B$ is in this set. Regarding the later event, $A(D_{0}) = \lambda_1(D_{0})$ implies $B_{0} = \lambda_0B_{0}$, so the $\lambda_1$ eigenvector of the $A$ used to compute $B$ is $D_{0}$. This demonstrates that all sign patterns for the components of $o(n)$ are equally likely. In other words, the distribution of the number of positive components in $o(n)$ is the binomial distribution $B(n, 1/2)$ and the fraction of positive components in $o(n)$ converges (in several senses) to $1/2$ as $n$ grows large.

Additionally, we can consider how $\lambda_1$ varies with $n$ in the case that $\mu = 0$ to determine when the blowup will occur. Füredi and Komlós (9) found for this case that $\lambda_1 < 2\sigma\sqrt{n} + O(n^{1/3} \log n)$ with probability tending to 1, so with probability tending to 1 the blow-up time shrinks to zero like $1/\sqrt{n}$, an order of $\sqrt{n}$ slower than in the $\mu > 0$ case.

**Case 3: $\mu < 0$.** For this final case, Füredi and Komlós (9) found that $\lambda_1 > 2\sigma\sqrt{n} - o(\sqrt{n})$ in probability tending to 1. The semicircle law gives a lower bound: $\lambda_1 > 2\sigma\sqrt{n} + o(\sqrt{n})$ in probability. So the blow-up time goes to zero like $1/\sqrt{n}$ in the unscaled system.

Note also that if we define a matrix $C = -A$, where $A$ is now the initial matrix $X(0)$ of Case 3, then $C$ satisfies the condition of
Case 1, \( \mu > 0 \). Thus the distance between the top eigenvector of \( C \) and \((1,\ldots,1)\sqrt{n}/n\) declines to zero in probability just as in Case 1. Furthermore, every other eigenvector of \( C \) is orthogonal to the largest one. Hence if \( \sigma < |\mu|/2 \), then with probability tending to 1, every other eigenvector acquires a mixture of positive and negative components in the large-\( n \) limit, including the bottom eigenvector of \( C \), which is the top eigenvector of \( A \). This establishes that in the case that \( \mu < 0 < |\mu|/2 \), the system ends up in a state with two factions with probability converging to 1 for all finite \( n \).

Numerical simulations of the case that \( \mu < 0 \) suggest the conjecture that the two factions are approximately equal in size for large \( n \). Furthermore, the derivation of Eq. 9 is in fact valid for all \( \mu \), so each trajectory rapidly decays from \( x_{ij}(0) \) toward \( x_{ij}(0) - \mu \) on \([0, 1/K]\) (Fig. 1C). This transient decay appears to extend beyond \( t = 1/K \) in numerical simulations. So, for example, if time is rescaled by \( 1/\sqrt{n} \) instead of \( 1/n \), we would hypothesize that (i) each trajectory makes a complete jump from \( x_{ij}(0) \) to \( x_{ij}(0) - \mu \) in the large-\( n \) limit and that (ii) from this point onward, the system behaves like an initial configuration of the \( \mu = 0 \) case and so separates into two equal factions en route to its blowup at \( 1/(2\sigma) \).

Discussion

In this final section, we review our results and their significance relative to previous work in structural balance theory. We then compare the predictions of the model with data, discuss potential criticisms of the model, and finish with some intriguing connections between the behavior of the model and recent social-psychological work on neutralizing two-sided conflicts.

Our first result is a demonstration that the model forms two factions in finite time across a broad set of initial conditions. As noted at the outset, similar demonstrations have not been possible for dynamic models of structural balance in earlier literature because these models contained so-called jammed states that could trap a social network before it reached a two-faction configuration (4, 5). The model of Kulakowski et al. by contrast has no such jammed states for generic initial conditions and hence provides a robust means for a social network to balance itself.

The second result of the paper is the discovery and characterization of a transition from global polarization to global harmony as the initial mean friendliness of the network crosses from nonpositive to positive values. Similar transitions have been observed in other models of structural balance but so far none has been characterized at a quantitative level. For example, Antal et al. (4) found a nonlinear transition from two cliques of equal size to a single unified clique as the fraction of positively signed edges at \( t = 0 \) was increased from 0 to 1 (see figure 5 of ref. 4). The authors provided a qualitative argument for this transition, but left open the problem of its quantitative detail. Our results both confirm the generality of their observations and provide a quantitative account of a transition analogous to theirs.

To complement the theoretical nature of our work and get a better sense of how the model behaves in practice, we can numerically integrate it for several cases of empirical social network data where the real-life outcomes of the time evolution are known. Our first example is based on a study by Zachary (12) who witnessed the breakup of a karate club into two smaller clubs. Prior to the separation, Zachary collected counts of the number of social contexts in which each pair of individuals interacted outside of the karate club, with the idea being that the more social contexts they shared, the greater the likelihood for information exchange. These counts, or capacities as Zachary called them, can be converted to estimates of friendliness and rivalry in many different ways. For a large class of such conversions, Eq. 1 predicts the same division that Zachary’s method found, which misclassified only 1 of the 34 club members (Fig. 2A and B).

A second example can be constructed from the data of a study by Axelrod and Bennett (13) regarding the aggregation of Allied and Axis powers during World War II. If we simply take the entries of their propensity \((i,j) \cdot \text{size}(i) \cdot \text{size}(j)\) matrix to be proportional to the friendliness felt between the various pairs of countries in the war, then running the model gives the correct Allied-Axis split for all countries except Denmark and Portugal (Fig. 2C).

Nevertheless, the model clearly contains several strong simplifications of the underlying social processes. The first of these is inherent to structural balance theory itself; it is a framework restricted to capture a particular kind of social situation, in which the need for consistency among one’s friendships and rivalries brings about the emergence of two factions. Extensions of the theory have considered models in which it is possible to have multiple mutual enemies and hence more than two factions (14), and also networks that are not complete graphs (3). However, our focus here has been on the basic theory, because as we have seen, obtaining a satisfactory dynamics even for this simplest form of structural balance has been an elusive challenge. Moreover, the
basic version of structural balance that we have considered here, with a complete graph of relationships and constraints leading to two factions, is relevant to a range of different situations. These span the kinds of settings discussed earlier in this section, including clubs, classrooms, and small organizations (15), as well as international relations during crisis (where a large set of nations can all mutually maintain friendly or unfriendly diplomatic relations) (13, 16).

Another consequence of the particular model studied here that has no direct analogue in real social situations is the divergence to infinity of the relationship strengths $x_{ij}$. However, because the purpose of the model is to study the pattern of signs that emerges, our main conclusions are based not on the actual magnitudes of these numbers but on the fact that the sign pattern eventually stabilizes at a point before the divergence. This stabilization of the sign pattern is our primary focus, and one could interpret the subsequent singularity as simply the straightforward realization of the sign pattern is our primary focus, and one could eventually stabilize at a point before the divergence. This stabilization itself can be viewed as taking place beyond the window of time over which the system corresponds to anything real. Alternately, one can imagine that as the community completes its separation into two groups, other social processes take over. For example, individuals with differing ideological views or social preferences may self-segregate, breaking the all-to-all assumption of the model. In other cases, mounting tensions may erupt into violence, reflecting a sort of bound on the relationship intensity achievable for pairs of nodes in the network.

Finally, we note that there is a large body of work in social psychology that studies issues such as the formation and reconciliation of factions from a much more empirical basis; see, for example, refs. 17 and 18. It is an interesting open problem to determine the extent to which the strictly mathematical development of the models here can be combined with the perspectives in this empirical body of literature, ultimately leading to a richer theory of these types of social processes.

ACKNOWLEDGMENTS. We thank Nick Trefethen and David Bindel for pointers to the literature on matrix Riccati equations. Research was supported in part by the John D. and Catherine T. MacArthur Foundation, a Google Research Grant, a Yahoo! Research Alliance Grant, an Alfred P. Sloan Foundation Fellowship, a Microsoft Research New Faculty Fellowship, a grant from the Air Force Office of Scientific Research, and National Science Foundation Grants CCF-0325453, BCS-0537606, CCF-0643934, IIS-0705774, and CISE-0835706.

Supporting Information

Marvel et al. 10.1073/pnas.1013213108

SI Text

Referenced Results. In the introduction of our paper, we assert that a variant of the dynamics proposed by Kulakovski et al. (1) generically balances an isolated triangle. We explain what we mean here.

**Theorem 1.** The system \( \dot{x}_{12} = x_{13}^2 x_{23}, \) \( \dot{x}_{13} = x_{12} x_{23}, \) \( \dot{x}_{23} = x_{12} x_{13} \) achieves balance when the initial values \( x_{12}(0), x_{13}(0), \) and \( x_{23}(0) \) are all unequal.

**Proof:** Multiplying each \( \dot{x}_i \) by \( x_i \) yields \( \dot{x}_{12} x_{12} = x_{13} x_{23} = x_{12} x_{23} x_{13}. \) Integrating these equalities gives the constraints \( x_{12}^2 - x_{13}^2 = C_1 \) and \( x_{12}^2 - x_{23}^2 = C_2, \) which partition the three-dimensional space of \( (x_{12}, x_{13}, x_{23}) \) into trajectories (with the direction of flow defined by the original dynamical system). Examination of this flow reveals that each initial condition \( (x_{12}(0), x_{13}(0), x_{23}(0)) \) with distinct coordinates flows into one of the four octants on which Heider balance holds; that is, where \( x_{12} \leq x_{13} = x_{23} > 0. \) Furthermore, these octants each act as separate trapping regions: Once a trajectory enters, it cannot leave. Hence, the theorem follows.

The next theorem regards the main system of the paper with a rescaling of time; \( X = n^{-1} X^2, \) where \( X \) is a real symmetric \( n \times n \) matrix. Recall that \( x_i(t) \) denotes the \( (i,j) \)th element of the solution matrix \( X(t) \) subject to the initial condition \( X(0) \) in the following. We abbreviate \( X(0) \) as \( A \) and \( x_i(0) \) as \( a_i. \) Suppose that the \( a_i, i \leq j, \) are drawn independently from distributions \( F_j \) with zero mean outside \([-K,K] \), and the off-diagonal distributions \( F_j \) have common expectation \( \mu \) and variance \( \sigma^2. \)

**Theorem 2.** \( \lim_{n \to \infty} x_{ij}(t) = a_{ij} - \mu + \mu / (1 - \mu) \) with probability 1 for \( t \in [0,1/K] \).

**Proof:** Regard each element of the limit \( n \to \infty \) as a selection and concatenation of \( \{a_{ij}\}_{i \leq j}, \{a_{ij}\}_{1 \leq i \leq n} \) to the elements \( \{a_{ij}\}_{i \leq j} \) selected in preceding steps. Now consider the partial sum of the Taylor series expansion of \( x_{ij}(t): \)

\[
x_{ij}(t) = \sum_{k=0}^{N} a_{ij} \frac{k!}{k!} t^k \quad \text{where} \quad a_{kn} = \frac{d^k x_{ij}}{dt^k} \bigg|_{t=0}. \quad [S1]
\]

The first step of the proof of Theorem 2 consists of proving that \( \lim_{n \to \infty} \lim_{m \to \infty} x_{ij}(t) \) converges to \( a_{ij} - \mu + \mu / (1 - \mu) \) with probability 1 on \([0,1/K]\) (see Lemma 1). The second step is to prove that \( \lim_{n \to \infty} \lim_{m \to \infty} x_{ij}(t) = \lim_{n \to \infty} x_{ij}(t) \) with probability 1 on \([1/K,1]\) (see Lemma 2). Because we can write \( \lim_{n \to \infty} x_{ij}(t) = \lim_{n \to \infty} x_{ij}^{(N)}(t) \), this amounts to showing that the two limits can be exchanged on \([0,1/K]\). The above theorem then follows trivially by a union bound.

**Lemma 1.** Under the assumptions of Theorem 2, \( \lim_{n \to \infty} \lim_{m \to \infty} x_{ij}(t) = a_{ij} - \mu + \mu / (1 - \mu) \) with probability 1 for \( t \in [0,1/K] \).

**Proof:** For the sake of generality, we present a proof with more mild assumptions than those of the rest of the paper: We require only that the moments of the \( F_j \) distributions be finite (and off-diagonal distributions to have mean \( \mu \), not that the \( a_{ij} \) values be bounded by \( K \) with probability 1. Define \( a_{ij} = \lim_{n \to \infty} a_{ij} \) (merely shorthand—we do not assume the limit exists). By a union bound, we have

\[
\Pr\left( \bigcap_{k=1}^{m} \{a_{ij} = \mu^{k+1}\} \right) \geq 1 - \sum_{k=1}^{m} \{1 - \Pr(a_{ij} = \mu^{k+1})\} \quad [S2]
\]

So if we can show that \( \Pr(a_{ij} = \mu^{k+1}) = 1 \) for all \( k \geq 1, \) then \( \Pr\left( \bigcap_{k=1}^{m} \{a_{ij} = \mu^{k+1}\} \right) = 1. \) In this case, \( \lim_{n \to \infty} \lim_{m \to \infty} x_{ij}(t) \) has the convergent Taylor series \( a_{ij} + \sum_{k=1}^{\infty} \mu^{k+1} \) on \([0,1/K]\) with probability 1, which proves the lemma.

Our task reduces to showing that \( \Pr(a_{ij} = \mu^{k+1}) = 1 \) for each \( k \geq 1. \) In order to do this, we need to compute the leading behavior of \( a_{ij} \) in \( n. \) To calculate the \( k \) time derivatives of \( x_{ij} \) in the formula for \( a_{ij} \) (see Eq. S1), we alternate between applying the chain rule of differential calculus and substituting in the right-hand side of \( x_{ij} = n^{-1} \sum x_{ik} x_{kj} \) (our system \( X = n^{-1} X^2 \) written in elementwise fashion). This gives

\[
a_{kn} = n^{-k} \sum a_{mi} a_{mi} a_{mj} \quad [S3]
\]

where the factor \( n^{-k} \) comes from the \( k \) factors of \( n^{-1} \) introduced by the \( k \) derivatives, and the factor \( 1/k! \) in the formula for \( a_{ij} \) cancels with a factor \( k! \) that arises from repeated applications of the chain rule. In Eq. S3, the dominant term is a sum of the edge value products of all simple length-(\( k+1 \)) paths between \( i \) and \( j. \) This sum contains \((n-2)!/(n-2-k)!\) terms. All other paths include fewer immediate nodes and thus have at least a factor of \( n \) fewer terms in their sums.

Our goal then for the remainder of the proof is to show that the first term of Eq. S3 is the only term that remains after taking \( n \) to infinity, and that it converges to \( \mu^{k+1} \) with probability 1. To simplify notation, let \( \ell \) denote the product of the edge values \( a_{ij} \) along a particular path of length \( k+1 \) (not necessarily simple) from \( i \) to \( j. \) and let \( L \) denote the set of all such products on paths with the same configuration, or pattern of connectivity. Denote the set of all \( L \) by \( \{L\}, \) and let \( S \) denote the one \( L \) in \( \{L\} \) consisting of simple paths of length \( k+1 \) from \( i \) to \( j. \)

Now observe that \( r_{ij} = \lim_{n \to \infty} n^{-k} \sum_{\ell \in \ell} \ell = 0 \) \( n^{-k} \sum_{\ell \in \ell} \ell = \mu^{k+1} \) \( \in \{a_{ij} = \mu^{k+1}\}. \) So by another union bound,

\[
\Pr(a_{ij} = \mu^{k+1}) \geq 1 - \Pr\left( \bigcup_{\ell \in \ell} \{\ell = \mu^{k+1}\} \right) - \sum_{\ell \in \ell} \Pr\left( \bigcup_{\ell \in \ell} \{\ell = \mu^{k+1}\} \right). \quad [S4]
\]

Hence, we are done if we can show that \( i \Pr(\lim_{n \to \infty} n^{-k} \sum_{\ell \in \ell} \ell = \mu^{k+1}) = 1 \) for \( S \) and \( ii \) \( \Pr(\lim_{n \to \infty} n^{-k} \sum_{\ell \in \ell} \ell = 0) = 1 \) for all other \( L. \) Although \( \sum_{\ell \in \ell} \ell \) is in general a sum of correlated random variables, it is possible to adapt a standard proof of the strong law of large numbers for uncorrelated random variables to prove both items. We do this next.

Let us prove \( i \) first. For brevity, let \( S_n = \sum_{\ell \in \ell} \ell \) and choose \( \nu \) to denote the number of nodes in the path configuration of \( L. \) For any positive \( c \) and \( r = 1.2. \ldots, \) Markov's inequality gives
Pr(\(|S_0| \geq (n e)^k\)) \leq E(|S_0|)/\((n e)^{2k}\). So if we can find an \(r\) such that \(E(|S_0|)/\((n e)^{2k}\) \leq C/n^2\) for some constant \(C\) (dependent on \(e\)), then \(\sum_n \Pr(S_n) \geq (n e)^k\) converges, and by the first Borel-Cantelli lemma, \(\Pr(n^{-k} \sum_{n \leq L} E(\epsilon_n^2) \geq e i.o.) = 0\) for all \(e > 0\) (where \(i.o.\) stands for infinitely often). Careful reflection reveals that \(\Pr(\epsilon_n^2) \geq e i.o.)\) (for, say, all rational \(e\)) is the complementary event of \(\lim_{n \to \infty} n^{-k} \sum_{n \leq L} E(\epsilon_n^2) = 0\) and so we have arrived at the desired result (ii).

Hence, in order to actually show (ii), we need to find an \(r\) such that \(E(|S_0|)/\((n e)^{2k}\) \leq C/n^2\). Consider \(\epsilon = 2: E(S_0^2) = \sum E(\epsilon_n^2)\), where each index of the sum ranges independently over \(L\). There are \((n - 2)!(n - 2 - v)\) paths \(\epsilon\) in \(L\), so there are fewer than \(n^{2v}\) terms in \(\sum E(\epsilon_n^2)\), where each range independently from 1 to \(n\). Because \(\epsilon < k\) for all \(L\) other than \(S\), we have \(E(|S_0|)/\((n e)^{2k}\) \leq D^{n^2}\) for some constant \(D\). Therefore, for any \(L\), \(\epsilon_n^2\) is given by \(\epsilon_n^2 = \epsilon_n^2\) and the proof of (ii) is complete.

Finally, to prove (i), start by replacing each factor \(a_{\epsilon_n}\) with \(b_{\epsilon_n} = a_{\epsilon_n}\) and then show how they together prove Lemma 2. Throughout the following, all statements hold with probability 1 unless stated otherwise.

As we found in the course of the proof of Lemma 1, the limits \(\lim_{n \to \infty} a_{\epsilon_n}\) exist for all \(k\) and are \(\mu^{k+1}\) on \([0, 1/\mu]\), so \(\lim_{n \to \infty} \sum_{k=0}^N a_{\epsilon_n}^k\) exists under the same conditions, and we call it \(x_{\epsilon_n}(t)\). This gives us the first ingredient: \(\lim_{n \to \infty} a_{\epsilon_n}(t) = x_{\epsilon_n}(t)\) for \(t \in [0, 1/\mu]\) and any \(N\). Additionally, from Lemma 1 we know that \(\lim_{n \to \infty} x_{\epsilon_n}(t)\) exists and is \(a_{\epsilon_n} = \mu^{1} < \mu/\mu(1 - \mu)\) on \([0, 1/\mu]\). We call this limit \(x_{\epsilon_n}(t)\) and write the second ingredient as \(\lim_{n \to \infty} x_{\epsilon_n}(t) = x_{\epsilon_n}(t)\) for \(t \in [0, 1/\mu]\).

Finally, as we saw in the proof of Lemma 1, \(\alpha_{\epsilon_n} = n^{-k} \sum_{i,j} a_{\epsilon_n i} a_{\epsilon_n ij} \cdots a_{\epsilon_n ij}\) (by definition, not just with probability 1), where the \(k\) indices \(m_j\) range independently from 1 to \(k\). Because each \(|a_{\epsilon_n}| < k\), we must have that \(|a_{\epsilon_n}| < k^{1+1}\), which implies \(x_{\epsilon_n}(t)\) converges uniformly to \(x_{\epsilon_n}(t)\). Therefore, \(\lim_{n \to \infty} x_{\epsilon_n}(t)\) converges uniformly to \(x_{\epsilon_n}(t)\): (iii) For every \(e > 0\), there exists an \(N_1\) such that for all \(N \geq N_1\), and all \(n\), \(\lim_{n \to \infty} x_{\epsilon_n}(t) - x_{\epsilon_n}(t) < \epsilon\).

To complete the proof of Lemma 2, we need to show that \(\lim_{n \to \infty} x_{\epsilon_n}(t)\) exists and is just \(x_{\epsilon_n}(t)\) on \([0, 1/\mu]\). Start by picking an \(e > 0\). Then by (iii), there exists an \(N_1\) such that if \(N > N_1\) then \(x_{\epsilon_n}(t) - x_{\epsilon_n}(t) < \epsilon\) for all \(n\). Similarly, (ii) implies that there exists an \(N_2\) such that if \(N > N_2\) then \(x_{\epsilon_n}(t) - x_{\epsilon_n}(t) < \epsilon\) for all \(n\). Finally, let \(N_3 = \max(N_1, N_2)\). Then by (i), we may choose an \(n_1\) such that if \(n > n_1\) then \(x_{\epsilon_n}(t) - x_{\epsilon_n}(t) < \epsilon\). Now define the following events:

\[ E_1 = [x_{\epsilon_n}(t) - x_{\epsilon_n}(t) < \epsilon]\]
\[ E_2 = [x_{\epsilon_n}(t) - x_{\epsilon_n}(t) < \epsilon]\]
\[ E_3 = [x_{\epsilon_n}(t) - x_{\epsilon_n}(t) < \epsilon]\]
\[ E_4 = [x_{\epsilon_n}(t) - x_{\epsilon_n}(t) < 3\epsilon]\]  

Observe that, in similar form to Eq. 4, \(E_1 \cap E_2 \cap E_3 \subset E_4\), so \(\Pr(E_4) \geq \Pr(E_1 \cap E_2 \cap E_3) = 1 - \Pr(E_1') \cap \Pr(E_2') \cap \Pr(E_3') = 1\) for all \(n \geq n_1\). Thus, \(\lim_{n \to \infty} x_{\epsilon_n}(t) - x_{\epsilon_n}(t) < 3\epsilon\) for all \(n \geq n_1\) and \(t \in [0, 1/\mu]\).

Unreferenced Results. In the main paper, we establish that \(\lambda_1 > 0\) in probability by way of Wigner’s semicircle law. However, we can also show that \(\lambda_1 > 0\) with high probability under another set of assumptions about how \(A\) is selected. Loosely speaking, the selection of the diagonal entries is more constrained in this alternative approach, whereas the selection of the off-diagonal entries is less so.

Suppose that all \(a_{ij}\), \(i \neq j = 1\), are chosen from the same distribution \(F\) and all \(a_{ij}\) are chosen from the same distribution \(G\). All selections are independent for \(i \leq j\). In addition, assume that the density of \(a_{ij} = a_{ij}\) is not entirely confined to negative values. Then we have the following theorem:

**Theorem 3.** \(\Pr(\lambda_1 \leq 0)\) is exponentially small in \(n\).

**Proof:** If \(\lambda_1 \leq 0\), then the corresponding matrix \(A\) is negative semidefinite. Therefore \(v^T A v \leq 0\) for every vector \(v\) (\(T\) denotes transposition). Let \(v_k\) denote the vector with \(+1\) and \(-1\) in its \((2k - 1)\)th and \(2k\)th components, respectively, and \(0\) in all other components. Then the event that \(v_k^T A v_k \leq 0\) is equivalent to the event that \(a_{2k-1} a_{2k} - a_{2k-1} a_{2k} a_{2k-1} < 0\). Note that the left-hand side of this final inequality has at least a constant probability of being positive. Now as \(k\) ranges from 1 to \(n/2\), we encounter \(n/2\) independent events, each having at least a constant probability of failure. Hence the probability that \(A\) is negative semidefinite is exponentially small in \(n\).
Very roughly, the final result of this supporting text says that in the case of negative \(\mu\), the distribution of the sum of the components of \(\omega_1 = (v_1, \ldots, v_n)\) retains no more than constant width in the large-\(n\) limit. So, for example, the mean of the components of \(\omega_1\) must shrink to zero at least as fast as \(1/n\). (Note that this convergence is faster than the \(1/\sqrt{n}\) convergence of means for independent and identically distributed random variables with finite mean and variance.) This result is consistent with the picture that when \(\mu\) is negative, the system is destined for two-sided conflict in the large-\(n\) limit.

To make the proof less cumbersome, suppose that the diagonal \(F_{ii}\) have common expectation \(\mu\) and variance \(\sigma^2\) just like the off-diagonal \(F_{ij}\).

**Theorem 4.** If \(\mu < 0\), then \(\lim_{n \to \infty} \Pr(\sum_{i=1}^{n} v_i < \epsilon) = 1\) for all \(\delta, \epsilon > 0\).

**Proof:** Consider the eigenvalue equation \(\sum_{j=1}^{n} a_j v_j = \lambda_i v_i\). Sum both sides over \(i\) and then rearrange and switch index labels to obtain

\[
\sum_{i=1}^{n} v_i \sum_{j=1}^{n} a_{ij} = \lambda_i \sum_{i=1}^{n} v_i. \tag{S6}
\]

We anticipate from standard results of random matrix theory that the left side of Eq. S6 is asymptotic to \(\mu n \sum_{i=1}^{n} v_i\) and the right side to \(2\sqrt{n} \sum_{i=1}^{n} v_i\). So let us define the two events

\[
E_1 = \left[ \sum_{i=1}^{n} v_i \sum_{j=1}^{n} a_{ij} - \frac{\mu}{n} \sum_{i=1}^{n} v_i \right] \leq \epsilon \]

\[
E_2 = \left[ \frac{\lambda_i}{n} \sum_{i=1}^{n} v_i - 2 \sqrt{n} \sum_{i=1}^{n} v_i \right] \leq \epsilon \tag{S7}
\]

where \([\cdot]\) denotes an event. Now, \(\Pr(E_1 \cap E_2) = 1 - \Pr(E'_1 \cup E'_2) \geq 1 - \Pr(E'_1) - \Pr(E'_2)\), where we have written \(E_i\) for the complementary event of \(E_i\). So if we can show that \(\lim_{n \to \infty} \Pr(E_1) = 1\) and \(\lim_{n \to \infty} \Pr(E_2) = 1\) for all \(\delta, \epsilon > 0\), it follows that \(\lim_{n \to \infty} \Pr(|\mu n - \sum_{i=1}^{n} v_i| < \epsilon + \epsilon n^{-\delta}) = 1\), which implies Theorem 4.

Hence our task reduces to showing that \(\lim_{n \to \infty} \Pr(E_1) = \lim_{n \to \infty} \Pr(E_2) = 1\). First consider \(\Pr(E_1)\). By Jensen’s inequality (2) and normalization of \(\omega_1\), \((n^{-1} \sum_{i=1}^{n} |v_i|^2) \leq n^{-1} \sum_{i=1}^{n} v_i^2 = n^{-1}\), so

\[
\sum_{i=1}^{n} |v_i| \leq \sqrt{n}. \tag{S8}
\]

Now define the additional event:

\[
E_3 = \left[ \max_{1 \leq i \leq n} \left| \sum_{j=1}^{n} (a_{ij} - \mu) \right| \leq \epsilon \right]. \tag{S9}
\]

By Eq. S8, we have \(E_3 \subset \max_{1 \leq i \leq n} \left| \sum_{j=1}^{n} (a_{ij} - \mu) \right| \leq \epsilon \subset [n^{-1/2} \sum_{i=1}^{n} |v_i|] \leq \epsilon \subset E_1\). The Bernstein inequality and a union bound together imply that \(\lim_{n \to \infty} \Pr(E_3) = 1\), because they give that

\[
\Pr(E'_3) \leq 2n \exp \left( -\frac{n^{3/2} \epsilon^2}{2\sigma^2 + (K + \mu) n^{-1/2} \sigma^3/3} \right). \tag{S10}
\]

So, \(\lim_{n \to \infty} \Pr(E_1) = 1\) as desired.

Regarding \(\Pr(E_2)\), we know that \(\lambda_i \in 2\sqrt{n} + o(\sqrt{n})\) in probability (3), which implies that \(\lim_{n \to \infty} \Pr(|\lambda_i/\sqrt{n} - 2\sigma| < \epsilon) = 1\) and therefore \(\lim_{n \to \infty} \Pr(E_2) = 1\) for all \(\epsilon > 0\) by Eq. S8. This completes the proof.

**Additional References.** Examples of two-sided social conflicts appear in a wide range of settings, including political party coalitions (4–8), interstate wars (9, 10), corporate standard setting (11), intertribal feuding (12), and even experiments in which group membership is based on arbitrary criteria (13–17).