Mixing patterns in networks

M. E. J. Newman

Department of Physics, University of Michigan, Ann Arbor, Michigan 48109-1120
and Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501
(Received 26 September 2002; published 27 February 2003)

We study assortative mixing in networks, the tendency for vertices in networks to be connected to other vertices that are like (or unlike) them in some way. We consider mixing according to discrete characteristics such as language or race in social networks and scalar characteristics such as age. As a special example of the latter we consider mixing according to vertex degree, i.e., according to the number of connections vertices have to other vertices: do gregarious people tend to associate with other gregarious people? We propose a number of measures of assortative mixing appropriate to the various mixing types, and apply them to a variety of real-world networks, showing that assortative mixing is a pervasive phenomenon found in many networks. We also propose several models of assortatively mixed networks, both analytic ones based on generating function methods, and numerical ones based on Monte Carlo graph generation techniques. We use these models to probe the properties of networks as their level of assortativity is varied. In the particular case of mixing by degree, we find strong variation with assortativity in the connectivity of the network and in the resilience of the network to the removal of vertices.

I. INTRODUCTION

The techniques of statistical physics were developed to study the properties of systems of many interacting particles, atoms, or molecules, but their applicability is wider than this, and recent work has fruitfully applied these techniques to economies, ecosystems, social interactions, the Internet, and many other systems of current interest. The component parts of these systems, the analogs of atoms and molecules, such things as traders in a market, or computers on the Internet, are not usually connected together on a regular lattice as the atoms of a crystal are. Nor indeed do their patterns of connection normally fit any simple low-dimensional structure. Instead they fall on some more generalized "network," which may be more or less random depending on the nature of the system. The broadening of the scope of statistical physics to cover these systems has therefore led us to the consideration of the structure and function of networks, as one of the fundamental steps to understanding real-world phenomena of many kinds. Useful reviews of work in this area can be found in Refs. [1–3].

Recent studies of network structure have concentrated on a small number of properties that appear to be common to many networks and can be expected to affect the functioning of networked systems in a fundamental way. Among these, perhaps the best studied are the "small-world effect" [4,5], network transitivity or "clustering" [5], and degree distributions [6,7]. Many other properties, however, have been examined and may be equally important, at least in some systems. Examples include resilience to the deletion of network nodes [8–12], navigability or searchability of networks [13–15], community structure [16–18], and spectral properties [19–21]. In this paper we study another important network feature, the correlations between properties of adjacent network nodes known in the ecology and epidemiology literature as "assortative mixing."

The very simplest representation of a network is a collection of points, usually called vertices or nodes, joined together in pairs by lines, usually called edges or links. More sophisticated network models may introduce other properties of the vertices or the edges. Edges for example may be directed—they point in one particular direction—or may have weights, lengths, or strengths. Vertices can also have weights or other numerical quantities associated with them, or may be drawn from some discrete set of vertex types. In the study of social networks, the patterns of connections between people in a society, it has long been known that edges do not connect vertices regardless of their property or type. Patterns of friendship between individuals for example are strongly affected by the language, race, and age of the individuals in question, among other things. If people prefer to associate with others who are like them, we say that the network shows assortative mixing or assortative matching. If they prefer to associate with those who are different it shows disassortative mixing. Friendship is usually found to be assortative by most characteristics.

Assortative mixing can have a profound effect on the structural properties of a network. For example, assortative mixing of a network by a discrete characteristic will tend to break the network up into separate communities. If people prefer to be friends with others who speak their own language, for example, then one might expect countries with more than one language to separate into communities by language. Assortative mixing by age could cause stratification of societies along age lines. And while the main focus of this paper is on social networks, it is reasonable to suppose that similar mixing effects are seen in nonsocial networks also. We will give some examples of this in Sec. III A.

In this paper, we study assortative mixing of various types using empirical network data, analytic models, and numerical simulation. We demonstrate that assortative (or disassortative) mixing is indeed present in many networks, show how it can be measured, and examine its effect on network structure and behavior. The outline of the paper is as follows.

TABLE I. The mixing matrix e_{ij} and the values of a_i and b_i for sexual partnerships in the study of Catania *et al.* [24]. After Morris [25].

| Men | Women | | | | |
---	Black	Hispanic	White	Other	a_i
Black	0.258	0.016	0.035	0.013	0.323
Hispanic	0.012	0.157	0.058	0.019	0.247
White	0.013	0.023	0.306	0.035	0.377
Other	0.005	0.007	0.024	0.016	0.053
b_i	0.289	0.204	0.423	0.084	

In Sec. II we study the effects of assortative mixing by discrete characteristics such as language or race. In Sec. III we study mixing by scalar properties such as age and particularly vertex degree; since degree is itself a property of the network topology, the latter type of mixing leads to some novel network structures not seen with other types. In Sec. IV we give our conclusions. A preliminary report of some of the results in this paper has appeared previously as Ref. [22].

II. DISCRETE CHARACTERISTICS

In this section we consider assortative mixing according to discrete or enumerative vertex characteristics. Such mixing can be characterized by a quantity e_{ij}, which we define to be the fraction of edges in a network that connect a vertex of type i to one of type j. On an undirected network this quantity is symmetric in its indices $e_{ij}=e_{ji}$, while on directed networks or bipartite networks it may be asymmetric. It satisfies the sum rules

$$\sum_{ij} e_{ij}=1, \quad \sum_{j} e_{ij}=a_i, \quad \sum_{i} e_{ij}=b_j, \quad (1)$$

where a_i and b_i are the fraction of each type of end of an edge that is attached to vertices of type i. On undirected graphs, where the ends of edges are all of the same type, $a_i=b_i$ [23].

For example, Table I shows data on the values of e_{ij} for mixing by race among sexual partners in a 1992 study carried out in the city of San Francisco, California [24]. This part of the study focused on heterosexuals, so this is a bipartite network, the two vertex types representing men and women, with edges running only between vertices of unlike types. This means that in this case the ends of an edge are different and the matrix e_{ij} is asymmetric. As the table shows, mixing is highly assortative in this network, with individuals strongly preferring partners from the same group as themselves.

A. Measuring discrete assortative mixing

To quantify the level of assortative mixing in a network we define an *assortativity coefficient* thus

$$r=\frac{\sum_i e_{ii}-\sum_i a_i b_i}{1-\sum_i a_i b_i}=\frac{\mathrm{Tr}\,\mathbf{e}-||\mathbf{e}^2||}{1-||\mathbf{e}^2||}, \quad (2)$$

where $\mathbf{e}$ is the matrix whose elements are e_{ij} and $||\mathbf{x}||$ means the sum of all elements of the matrix $\mathbf{x}$. This formula gives $r=0$ when there is no assortative mixing, since $e_{ij}=a_i b_j$ in that case, and $r=1$ when there is perfect assortative mixing and $\Sigma_i e_{ii}=1$. If the network is perfectly disassortative, i.e., every edge connects two vertices of different types, then r is negative and has the value

$$r_{\min}=-\frac{\sum_i a_i b_i}{1-\sum_i a_i b_i}, \quad (3)$$

which lies in general in the range $-1 \leq r < 0$. One might ask what this value signifies. Why do we not simply have $r=-1$ for a perfectly disassortative network? The answer is that a perfectly disassortative network is normally closer to a randomly mixed network than is a perfectly assortative network. When there are several different vertex types (e.g., four in the case shown in Table I) then random mixing will most often pair unlike vertices, so that the network appears to be mostly disassortative. Therefore, it is appropriate that the value $r=0$ for the random network should be closer to that for the perfectly disassortative network than for the perfectly assortative one.

A quantity with properties similar to those of Eq. (2) has been proposed previously by Gupta, Anderson, and May [26]. However, the definition of Gupta, Anderson, and May gives misleading results in certain situations, such as, for example, when one type of vertex is much less numerous than other types, as is the case in Table I. In this paper therefore we use Eq. (2), which does not suffer from this problem. The difference between the two measures is discussed in more detail in Appendix A.

Using the values from Table I in Eq. (2), we find that $r=0.621$ for the network of sexual partnerships, implying, as we observed already, that this network is strongly assortative by race—individuals draw their partners from their own group far more often than one would expect on the basis of pure chance.

As another example of the application of Eq. (2), consider the network studied by Girvan and Newman [16] representing the timetable of American college football games, in which vertices represent universities and colleges, and edges represent regular season games between teams during the year in question. Colleges are grouped into "conferences," which are defined primarily by geography, and teams normally play more often against other teams in their own conference than they do against teams from other conferences. In other words, there should be assortative mixing of colleges by conference in the schedule network. For the 2000 season schedule studied in Ref. [16], we find a value for the

assortativity coefficient of $r=0.586$, again indicating strong assortative mixing, i.e., colleges do indeed play games with their conference partners to a substantially greater degree than one would expect in a randomly mixed network.

It is also useful to know the expected statistical error on the value of r, so that we can evaluate the significance of our results. One way to calculate this error is to use the jackknife method [27]. Regarding each of the M edges in a network as an independent measurement of the contributions to the elements of the matrix **e**, we can show that the expected standard deviation σ_r on the value of r satisfies

$$\sigma_r^2 = \sum_{i=1}^{M} (r_i - r)^2, \qquad (4)$$

where r_i is the value of r for the network in which the ith edge is removed. For example, in the case of the matrix of Table I this gives $\sigma_r=0.014$, which, when compared with the value $r=0.621$ shows that our finding of assortative mixing is strongly statistically significant—a 40σ result.

Although it has a rather different physical interpretation, the coefficient r is mathematically similar to the intraclass correlation coefficients used in statistics to compare measurements across different groups in a population [28]. Standard results for errors on intraclass correlations can be adapted to the present case to show that another estimate of the error on r is [29]

$$\sigma_r^2 = \frac{1}{M} \frac{\sum_i a_i b_i + \left[\sum_i a_i b_i \right]^2 - \sum_i a_i^2 b_i - \sum_i a_i b_i^2}{1 - \sum_i a_i b_i}, \quad (5)$$

which gives $\sigma_r=0.012$ for the data of Table I, comparable with the jackknife method. Either method for estimating σ_r will be adequate for most purposes—the choice between them is a matter of convenience.

B. Models of discrete assortative networks

The generalized random graph models of networks studied in the past by various authors [30–36] can be extended to the case of assortative mixing on discrete characteristics. Suppose we are told the degree distribution $p_k^{(i)}$ for vertices of type $i=1,\ldots,n$ in a network and the value of the mixing matrix e_{ij}. Implicitly, we are also told the values of the quantities a_i and b_i, since we can extract them from e_{ij} using Eq. (1). We consider the ensemble of all graphs with these values of $p_k^{(i)}$ and e_{ij}, which gives us a random graph model similar in spirit to that of Refs. [30–36] for the case of specified degree distribution only. Many properties of this ensemble can be calculated exactly in the limit of large system size, as we now demonstrate.

Suppose that a vertex of type i has degree k. The k vertices at the other ends of the edges attached to this vertex are divided among the n possible vertex types according to some partition $\{r_1, r_2, \ldots, r_n\}$, where $\Sigma_j r_j = k$. The probability

that the partition $\{r_j\}$ takes a particular value is given by the multinomial distribution

$$P^{(i)}(k,\{r_j\}) = k! \prod_j \frac{1}{r_j!} \left[\frac{e_{ij}}{\sum_j e_{ij}} \right]^{r_j}. \qquad (6)$$

Now, generalizing Ref. [35], we define a generating function for the distributions of the numbers of edges connecting to each type of vertex:

$$G_0^{(i)}(x_1, x_2, \ldots, x_n) = \sum_{k=0}^{\infty} p_k^{(i)} \sum_{\{r_j\}} \delta\left(k, \sum_j r_j \right)$$
$$\times P^{(i)}(k,\{r_j\}) x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}. \quad (7)$$

Performing the sum over $\{r_j\}$, this gives

$$G_0^{(i)}(x_1, x_2, \ldots, x_n) = \sum_k p_k^{(i)} \left[\frac{\sum_j e_{ij} x_j}{\sum_j e_{ij}} \right]^k$$
$$= G_0^{(i)} \left(\frac{\sum_j e_{ij} x_j}{\sum_j e_{ij}} \right), \qquad (8)$$

where

$$G_0^{(i)}(x) = \sum_k p_k^{(i)} x^k \qquad (9)$$

is the fundamental generating function for the degree distribution $p_k^{(i)}$, as defined in Ref. [35]. Similarly, for the edges connected to a vertex of type i reached by following a randomly chosen edge on the graph we have

$$G_1^{(i)}(x_1, x_2, \ldots, x_n) = G_1^{(i)} \left(\frac{\sum_j e_{ij} x_j}{\sum_j e_{ij}} \right) \qquad (10)$$

with

$$G_1^{(i)}(x) = \frac{\sum_k k p_k^{(i)} x^{k-1}}{\sum_k k p_k^{(i)}} = \frac{1}{z_i} \frac{dG_0^{(i)}}{dx} \bigg|_{x=1}, \qquad (11)$$

where $z_i \equiv G_0^{(i)\prime}(1)$ is the mean degree for type i vertices.

Now we consider the total number of vertices reachable by following an edge that arrives at a vertex of type i. This number has a distribution that is generated by a generating function $H_1^{(i)}$ satisfying a self-consistency condition of the form

$$H_1^{(i)}(x) = xG_1^{(i)}[H_1^{(1)}(x), \ldots, H_1^{(n)}(x)]. \qquad (12)$$

And similarly the distribution of the number of vertices reachable from a randomly chosen vertex of type i is generated by

$$H_0^{(i)}(x) = xG_0^{(i)}[H_1^{(1)}(x), \ldots, H_1^{(n)}(x)]. \qquad (13)$$

By solving these two equations simultaneously, we can derive the complete component size distribution in our model. Here however, we will just calculate some of the more important average statistics of our networks from the generating functions. For example, the mean number s_i of vertices reachable from a vertex of type i is

$$s_i = \frac{dH_0^{(i)}}{dx}\bigg|_{x=1} = 1 + G_0^{(i)\prime}(1)\frac{\sum_j e_{ij}H_1^{(j)\prime}(1)}{\sum_j e_{ij}}. \qquad (14)$$

We can write this in matrix form thus as

$$\mathbf{s} = \mathbf{1} + \mathbf{m}_0 \cdot \mathbf{H}_1'(1), \qquad (15)$$

where $\mathbf{m}_0$ is the matrix with elements $z_i e_{ij}/a_i$. Differentiating Eq. (12), we then find that

$$\mathbf{s} = \mathbf{1} + \mathbf{m}_0 \cdot [\mathbf{I} - \mathbf{m}_1]^{-1} \cdot \mathbf{1}, \qquad (16)$$

where $\mathbf{1}$ is the vector whose elements are all 1, and the matrix $\mathbf{m}_1$ has elements

$$[\mathbf{m}_1]_{ij} = \frac{z_2^{(i)}}{z_1^{(i)}} \frac{e_{ij}}{a_i} \qquad (17)$$

with $z_1^{(i)} \equiv z_i$ being the mean degree of type i vertices and $z_2^{(i)}$ being the mean number of neighbors at distance two from a type i vertex.

When the density of edges on the graph is small, $\mathbf{s}$ in Eq. (16) is finite, but it grows with increasing density and then diverges when the determinant of the matrix $\mathbf{I} - \mathbf{m}_1$ reaches its first zero. This point marks the phase transition at which a giant component first appears in the network, similar to the phase transition seen in uncorrelated random graphs. The condition $\det(\mathbf{I} - \mathbf{m}_1) = 0$ for the position of the phase transition is the equivalent of the condition of Molloy and Reed [32] for the position of the phase transition in an uncorrelated random graph of arbitrary degree distribution.

The size of the giant component can also be calculated in a straightforward manner. We define u_i to be the probability that a vertex of type i, reached by following a randomly chosen edge in the graph, does not belong to the giant component. This probability is precisely equal to the probability that none of the neighbors of that vertex are themselves members of the giant component, and hence u_i satisfies the self-consistency condition

$$u_i = G_1^{(i)}(u_1, \ldots, u_n). \qquad (18)$$

The probability that a randomly chosen vertex of type i is not a member of the giant component is then $G_0^{(i)}(u_1, \ldots, u_n)$, and the overall fraction S of vertices that are in the giant component is given by

$$S = 1 - \sum_i \frac{a_i}{z_i} G_0^{(i)}(u_1, \ldots, u_n), \qquad (19)$$

where we have made use of the fact that the fraction of vertices of type i in the network is equal to a_i/z_i.

The simultaneous solution of Eqs. (18) and (19) gives us the value of S for our network. In many cases we find that these equations are not solvable in closed form, but they can easily be solved numerically by iteration of Eq. (18) from suitable starting values for the u_i, and then substituting the result into Eq. (19).

C. Simulating discrete assortative networks

We would also like to be able to generate random networks with a given level of assortative mixing, in order to check our analytical results and also for use as substrates for other models, such as, for example, epidemiological models. A simple algorithm for achieving this is the following.

(1) We first choose a size for our graph in terms of the number M of edges and we draw M edges from the distribution e_{ij}. That is, we generate M edges, each one identified by the types of the vertices that it connects, in some manner such that the fraction of edges connecting vertices of types i and j tends to e_{ij} as M becomes large. In practice, a simple transformation method works well [37].

(2) We count the number of ends of edges of each type i, to give the sums m_i of the degrees of vertices in each class. We calculate the expected number n_i of vertices of each type from $n_i = m_i/z_i$ (rounded to the nearest integer), where z_i is the desired mean degree of vertices of type i.

(3) We draw n_i vertices from the desired degree distribution $p_k^{(i)}$ for type i. In general the degrees of these vertices will not sum exactly to m_i as we want them to. So we choose one vertex at random, discard it, and draw another from the distribution $p_k^{(i)}$ until the sum does equal m_i.

(4) We pair up the m_i ends of edges of type i at random with the vertices we have generated, so that each vertex has the number of attached edges corresponding to its chosen degree.

(5) We repeat from step (3) for each vertex type.
This method correctly generates assortatively mixed graphs with the given e_{ij} in the limit of large graph size. In Sec. III C of this paper we give some examples of simulations of assortatively mixed networks.

III. ASSORTATIVE MIXING BY SCALAR PROPERTIES

A similar, but distinct, form of assortative mixing is mixing that depends on one or more scalar properties of network vertices. A classic example of mixing of this type seen in many social networks is assortative mixing by age. In Fig. 1 (top panel) we show a scatter plot of the ages of marriage partners in the 1995 U.S. National Survey of Family Growth
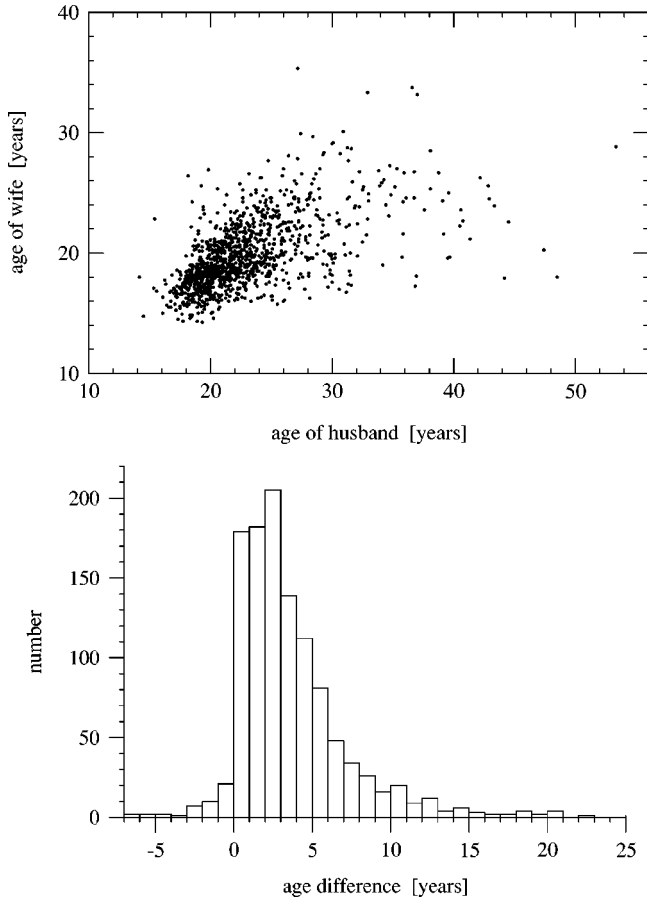
FIG. 1. Top, scatter plot of the ages of 1141 married couples at time of marriage, from the 1995 U.S. National Survey of Family Growth [38]. Bottom, a histogram of the age differences (male minus female) for the same data.

[38]. As is clear from the figure, there is a strong positive correlation between the ages, with most of the density in the distribution lying along a rough diagonal in the plot; people, it appears, prefer to marry others of about the same age, although there is some bias towards husbands being older than their wives. In the bottom panel of the same figure we show a histogram of the age differences in the study, which emphasizes the same conclusion [39].

By analogy with the developments of Sec. II, we can define a quantity e_{xy}, which is the fraction of all edges in the network that join together vertices with values x and y for the age or other scalar variable of interest. The values x and y might be either discrete in nature (e.g., integers, such as age to the nearest year) or continuous (exact age), making e_{xy} either a matrix or a function of two continuous variables. Here, for simplicity, we concentrate on the discrete case, but generalization to the continuous case is straightforward.

As before, we can use the matrix e_{xy} to define a measure of assortativity. We first note that e_{xy} satisfies the sum rules

$$\sum_{xy} e_{xy}=1, \quad \sum_{y} e_{xy}=a_x, \quad \sum_{x} e_{xy}=b_y, \qquad (20)$$

where a_x and b_y are, respectively, the fraction of edges that start and end at vertices with values x and y. (On an undi-

rected, unipartite graph, $a_x=b_x$.) Then, if there is no assortative mixing $e_{xy}=a_xb_y$. If there is assortative mixing one can measure it by calculating the standard Pearson correlation coefficient thus

$$r=\frac{\sum_{xy} xy(e_{xy}-a_xb_y)}{\sigma_a\sigma_b}, \qquad (21)$$

where σ_a and σ_b are the standard deviations of the distributions a_x and b_y. The value of r lies in the range $-1\leqslant r \leqslant 1$, with $r=1$ indicating perfect assortativity and $r=-1$ indicating perfect disassortativity (i.e., perfect negative correlation between x and y). For the age data from Fig. 1, for example, we find that $r=0.574$, indicating strong assortative mixing once more.

One can construct in a straightforward manner a random graph model of a network with this type of mixing exactly analogous to the model presented in Sec. II B. It is also possible to generate random representative networks from the ensemble defined by e_{xy} using the algorithm described in Sec. II C. In this paper however, rather than working further on the general type of mixing described here, we will concentrate on one special example of assortative mixing by a scalar property, which is particularly important for many of the networks we are interested in, namely mixing by vertex degree.

A. Mixing by vertex degree

In general, scalar assortative mixing of the type described above requires that the vertices of the network of interest have suitable scalar properties attached to them, such as age or income in social networks. In many cases, however, data are not available for these properties to allow us to assess whether the network is assortatively mixed. But there is one scalar vertex property that is always available for every network, and that is vertex degree. So long as we know the network structure we always know the degree of a vertex, and then we can ask whether vertices of high-degree preferentially associate with other vertices of high degree. Do the gregarious people hang out with other gregarious people? This has been a topic of considerable discussion in the physics literature [40–44]. As we will show, many real-world networks do show significant assortative (or disassortative) mixing by vertex degree.

Assortative mixing by degree can be quantified in exactly the same way as for other scalar properties of vertices, using Eq. (21). Taking the example of an undirected network and using the notation of Ref. [22], we define e_{jk} to be the fraction of edges that connect vertices of degrees j and k. In fact, we choose j and k to be the *excess degrees* of the vertices (also called *remaining degree* in Ref. [22]), which are one less than the degrees of the vertices themselves. This is because in most cases we are interested in the number of edges attached to a vertex other than the particular edge we are at the moment looking at.

If the degree distribution of the graph as a whole is p_k, i.e., p_k is the probability that a randomly chosen vertex will

have degree k, then the excess degree of the vertex at the end of an edge is distributed according to [35]

$$q_k = \frac{(k+1)p_{k+1}}{z}, \qquad (22)$$

where $z = \Sigma_k k p_k$ is the mean degree in the network. The distribution q_k is related to e_{jk} via

$$\sum_j e_{jk} = q_k. \qquad (23)$$

The correct assortativity coefficient for mixing by vertex degree in an undirected network is

$$r = \frac{\sum_{jk} jk(e_{jk} - q_j q_k)}{\sigma_q^2}, \qquad (24)$$

where σ_q is the standard deviation of the distribution q_k. For a directed network the equivalent expression is

$$r = \frac{\sum_{jk} jk(e_{jk} - q_j^{\text{in}} q_k^{\text{out}})}{\sigma_{\text{in}} \sigma_{\text{out}}}, \qquad (25)$$

where e_{jk} is now the probability that a randomly chosen directed edge leads into a vertex of in-degree j and out of a vertex of out-degree k [45]. For the purposes of calculating r for an actual network of specified vertices and edges, we can rewrite this in the form

$$r = \frac{\sum_i j_i k_i - M^{-1} \sum_i j_i \sum_{i'} k_{i'}}{\sqrt{\left[\sum_i j_i^2 - M^{-1}\left(\sum_i j_i\right)^2\right]\left[\sum_i k_i^2 - M^{-1}\left(\sum_i k_i\right)^2\right]}}, \qquad (26)$$

where j_i and k_i are the excess in-degree and out-degree of the vertices that the ith edge leads into and out of respectively, and M is again the number of edges. For an undirected network we can use the same formula—we simply replace each undirected edge by two directed ones leading in opposite directions. Alternatively, one can apply directly the formula given in Ref. [22], Eq. (4), to the undirected network. As before, errors on the measured values of r can be calculated using the jackknife method and Eq. (4).

In Table II we show the measured values of r for degree correlations in undirected and directed networks of a variety of different types, along with the expected errors on these values. The table reveals an interesting feature: essentially all the social networks examined are significantly assortative by degree, i.e., high-degree vertices tend to be connected to other high-degree vertices, while all the technological and biological networks are disassortative. Three of the values for r, for the network of student relationships, the power grid, and the graph of software dependencies, are null results,

meaning that they are not statistically different from zero. All the others however fit the pattern clearly, with positive values of r for the social networks and negative values for all the others.

What is the explanation of this phenomenon? In all probability, there are a number of different mechanisms at work. Some possibilities are the following.

(1) In the social networks it is entirely possible, and is often assumed in the sociological literature, that similar people do attract one another, and therefore that there could be a real preference among gregarious people for association with other gregarious people, and similarly for hermits.

(2) On the other hand, the networks of collaborations between academics, actors, and businesspeople considered here are affiliation networks, i.e., networks in which people are connected together by membership of common groups (authors of a paper, actors in a film, etc.). Since all members of a group are thus connected to all other members, the positive correlations between degrees may at least in part reflect the fact that the members of a large (small) group are connected to the other members of the same large (small) group.

(3) In the Internet and the World Wide Web there may be organizational reasons for degree anticorrelation between vertices. The high-degree vertices in these networks are often connectivity providers (internet) or directories (web), which by definition tend to be connected to the "little people"—the individual service subscribers in the case of the internet or the individual web pages on the web.

(4) Maslov and Sneppen [41] have shown that disassortativity can be produced as a finite-size effect by the constraint that no two vertices in a network are connected by more than one edge. This constraint causes high-degree vertices to "repel" one another, producing negative values of r. This explanation could account for at least a part of the disassortative mixing we see in the Internet, the protein and metabolic networks, and the food webs, although it cannot be applied directly to the web and neural networks, for which vertex pairs can and often do have more than one connection.

It appears therefore that some of the degree correlations we see in our networks could have real social or organizational origins, while others may be artifacts of the types of networks we are looking at and the constraints that are placed on their structure.

B. Models of assortative mixing by degree

In Ref. [22] we studied the ensemble of graphs that have a specified value of the matrix e_{jk} and solved exactly for its average properties using generating function methods similar to those of Sec. II B. We showed that the phase transition at which a giant component first appears in such networks occurs at a point given by $\det(\mathbf{I} - \mathbf{m}) = 0$, where $\mathbf{m}$ is the matrix with elements $m_{jk} = ke_{jk}/q_j$. One can also calculate exactly the size of the giant component, and the distribution of sizes of the small components below the phase transition. While these developments are mathematically elegant, however, their usefulness is limited by the fact that the generating functions involved are rarely calculable in closed form for arbitrary specified e_{jk}, and the determinant of the matrix ($\mathbf{I}$

TABLE II. Size n, degree assortativity coefficient r, and expected error σ_r on the assortativity, for a number of social, technological, and biological networks, both directed and undirected. Social networks: coauthorship networks of (a) physicists and biologists [46] and (b) mathematicians [47], in which authors are connected if they have coauthored one or more articles in learned journals; (c) collaborations of film actors in which actors are connected if they have appeared together in one or more movies [5,7]; (d) directors of fortune 1000 companies for 1999, in which two directors are connected if they sit on the board of directors of the same company [48]; (e) romantic (not necessarily sexual) relationships between students at a U.S. high school [49]; (f) network of email address books of computer users on a large computer system, in which an edge from user A to user B indicates that B appears in A's address book [50]. Technological networks: (g) network of high voltage transmission lines in the Western States Power Grid of the United States [5]; (h) network of direct peering relationships between autonomous systems on the Internet, April 2001 [51]; (i) network of hyperlinks between pages in the World Wide Web domain nd.edu, *circa* 1999 [52]; (j) network of dependencies between software packages in the GNU/Linux operating system, in which an edge from package A to package B indicates that A relies on components of B for its operation. Biological networks: (k) protein-protein interaction network in the yeast *S. Cerevisiae* [53]; (l) metabolic network of the bacterium *E. Coli* [54]; (m) neural network of the nematode worm *C. Elegans* [5,55]; tropic interactions between species in the food webs of (n) Ythan Estuary, Scotland [56] and (o) Little Rock Lake, Wisconsin [57].

	Group	Network	Type	Size n	Assortativity r	Error σ_r
	a	Physics coauthorship	undirected	52 909	0.363	0.002
	a	Biology coauthorship	undirected	1 520 251	0.127	0.0004
	b	Mathematics coauthorship	undirected	253 339	0.120	0.002
Social	c	Film actor collaborations	undirected	449 913	0.208	0.0002
	d	Company directors	undirected	7 673	0.276	0.004
	e	Student relationships	undirected	573	−0.029	0.037
	f	Email address books	directed	16 881	0.092	0.004
	g	Power grid	undirected	4 941	−0.003	0.013
Technological	h	Internet	undirected	10 697	−0.189	0.002
	i	World Wide Web	directed	269 504	−0.067	0.0002
	j	Software dependencies	directed	3 162	−0.016	0.020
	k	Protein interactions	undirected	2 115	−0.156	0.010
	l	Metabolic network	undirected	765	−0.240	0.007
Biological	m	Neural network	directed	307	−0.226	0.016
	n	Marine food web	directed	134	−0.263	0.037
	o	Freshwater food web	directed	92	−0.326	0.031

−m) almost never is. In this paper, therefore, we take an alternative approach, making use of computer simulation.

We would like to generate on a computer a random network having, for instance, a particular value of the matrix e_{jk}. [This also fixes the degree distribution, via Eq. (23).] In Ref. [22] we discussed one possible way of doing this using an algorithm similar to that of Sec. II C. One would draw edges from the desired distribution e_{jk} and then join the degree k ends randomly in groups of k to create the network. (This algorithm has also been discussed recently by Dorogovtsev, Mendes, and Samukhin [42].) As we pointed out, however, this algorithm is flawed because in order to create a network without any dangling edges the number of degree k ends must be a multiple of k for all k. It is very unlikely that these constraints will be satisfied by chance, and there does not appear to be any simple way of arranging for them to be satisfied without introducing bias into the ensemble of graphs. Instead, therefore, we use a Monte Carlo sampling scheme which is essentially equivalent to the Metropolis–Hastings method widely used in the mathematical and social sciences for generating model networks

[58,59]. The algorithm is as follows.

(1) Given the desired edge distribution e_{jk}, we first calculate the corresponding distribution of excess degrees q_k from Eq. (23), and then invert Eq. (22) to find the degree distribution:

$$p_k = \frac{q_{k-1}/k}{\sum_j q_{j-1}/j}. \qquad (27)$$

Note that this equation cannot tell us how many vertices there are of degree zero in the network. This information is not contained in the edge distribution e_{jk} since no edges connect to degree-zero vertices, and so must be specified separately. On the other hand, most of the properties of networks with which we will be concerned here do not depend on the number of degree-zero vertices, so we can safely set $p_0 = 0$ for the purposes of this paper.

(2) We draw a degree sequence, a specific set k_i of degrees of the vertices $i = 1, \ldots, N$, from the distribution p_k,

and connect vertices together randomly in pairs to generate a random graph, as described, for instance, by Molloy and Reed [32].

(3) We choose two edges at random from the graph. Let us denote these by the vertex pairs (v_1, w_1) and (v_2, w_2) that they connect.

(4) We measure the excess degrees j_1, k_1, j_2, k_2 of the vertices v_1, w_1, v_2, w_2 and then we remove the two edges and replace them with two new ones (v_1, v_2) and (w_1, w_2) with probability

$$
P = \begin{cases} \dfrac{e_{j_1 j_2} e_{k_1 k_2}}{e_{j_1 k_1} e_{j_2 k_2}} & \text{if } e_{j_1 j_2} e_{k_1 k_2} < e_{j_1 k_1} e_{j_2 k_2} \\[3mm] 1 & \text{otherwise.} \end{cases} \tag{28}
$$

(5) Repeat from step (3).

Clearly this swap procedure preserves the degree sequence. It is also ergodic over the set of graphs with that degree sequence, i.e., it can reach any graph within that set in a finite number of moves. To see this, consider any configuration of the graph other than the desired target configuration and choose any vertex pair that is not joined by an edge in that configuration but *is* joined by an edge in the target. These vertices must necessarily each be attached to at least one other edge that does not exist in the target configuration. Take these edges and perform the swap procedure on them. This always increases the number of edges that the configuration has in common with the target. And since it is always possible to do this, it immediately follows that any target configuration can be reached in at most M such moves, where M is the number of edges in the network. (Actually, $M-1$ will suffice, since the last edge will always automatically be in the correct position by a process of elimination.)

Our algorithm also satisfies detailed balance. We would like to sample graph configurations μ with probabilities [60]

$$
p_\mu = \prod_{i=1}^{M} e_{j_i k_i}, \tag{29}
$$

where j_i, k_i are the excess degrees of the vertices at the ends of the ith edge. It is trivial to show that with the choice of transition probabilities given in Eq. (28), we satisfy detailed balance in the form

$$
p_\mu P(\mu \rightarrow \nu) = p_\nu P(\nu \rightarrow \mu) \tag{30}
$$

for all pairs μ, ν of states.

Since our algorithm satisfies both ergodicity and detailed balance, it immediately follows [37] that in the limit of long time it samples graph configurations correctly from the distribution Eq. (29). It also turns out to be a reasonably efficient algorithm in practice. In the simulations reported here, the mean fraction of proposed Monte Carlo moves that was accepted never fell below 50% for any parameter values.

To use this algorithm we also need to choose a value for the matrix e_{jk}. We have a lot of freedom about how we do this. Suppose for example that we wish to simulate an undi-

rected network, so that e_{jk} is symmetric. A rank-n symmetric matrix has $\frac{1}{2}n(n+1)$ degrees of freedom, n of which are fixed in this case by the requirement, Eq. (23), that the rows and columns sum to q_k, leaving $\frac{1}{2}n(n-1)$ that can be freely chosen. [One could think for example of choosing the $\frac{1}{2}n(n-1)$ off-diagonal elements of e_{jk} and then satisfying Eq. (23) by choosing the n diagonal elements to give the row and column sums the desired values.]

A simple example of a disassortative choice for e_{jk} is the set of matrices taking the form

$$
e_{jk}^{(d)} = q_j x_k + x_j q_k - x_j x_k, \tag{31}
$$

where x_k is any distribution normalized such that $\Sigma_k x_k = 1$. It is easy to see that this choice satisfies the constraints on e_{jk}, and the resulting value of r is

$$
r_d = -\frac{(\mu_q - \mu_x)^2}{\sigma_q^2}, \tag{32}
$$

which is always negative (or zero). Here μ_q and μ_x are the means of the distributions q_k and x_k, respectively.

Being a probability, e_{jk} is also constrained to lie in the range $0 \le e_{jk} \le 1$ for all j, k. To ensure that Eq. (31) never becomes negative we should choose x_k to decay faster than q_k.

Suppose for example that we are interested, as many people seem to be these days, in networks that have power-law degree distributions, $p_k \sim k^{-\tau}$ [1–3,7,34,52,61–65]. True power laws unfortunately are troublesome to deal with; the crucial distribution q_k has a divergent mean unless $\tau > 3$, which it seldom is for real-world networks (see, for instance, Ref. [2], Table II). Instead, therefore, following Ref. [35], we here examine the exponentially truncated power-law distribution

$$
p_k = \frac{k^{-\tau} e^{-k/\kappa}}{\text{Li}_\tau(e^{-1/\kappa})} \quad \text{for } k \ge 1, \tag{33}
$$

where the function $\text{Li}_n(x)$, which acts here as a normalizing constant, is the nth polylogarithm of x. This gives a similar distribution $q_k \sim (k+1)^{-\tau+1} e^{-(k+1)/\kappa}$ for the excess degree, and we choose x_k to have the same functional form, but with a different cutoff parameter κ', where $\kappa' < \kappa$, to ensure that $e_{jk} > 0$. In Fig. 2 we show a plot of the resulting e_{jk} for $\tau = 2.5$, $\kappa = 100$, $\kappa' = 10$. The disassortative nature of this choice for e_{jk} is evident from the concentration of probability along the edges of the matrix in the figure.

Introducing a cutoff in the degree distribution also provides us with a parameter, namely κ, that can, as we will shortly see, be conveniently manipulated to produce a phase transition at which a giant-component appears in the network.

We can also make an assortative matrix e_{jk} by writing

$$
e_{jk}^{(a)} = 2q_j q_k - e_{jk}^{(d)}, \tag{34}
$$

which gives a value for the assortativity coefficient of $r_a = -r_d$, while still having the same degree distribution. More
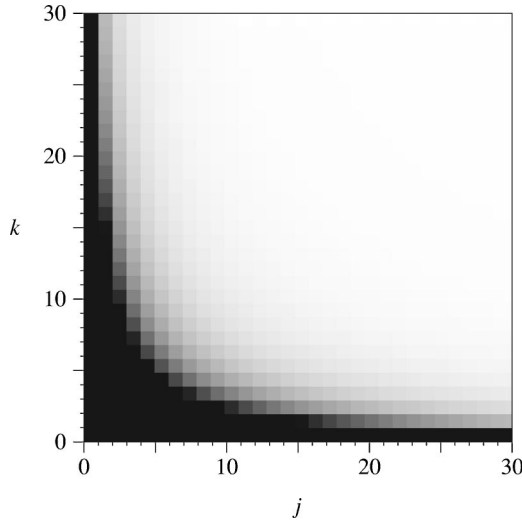
FIG. 2. Density plot of the matrix $e_{jk}^{(d)}$, Eq. (31), for $\tau = 2.5$, $\kappa = 100$, $\kappa' = 10$. Darker squares represent higher values of e_{jk}.
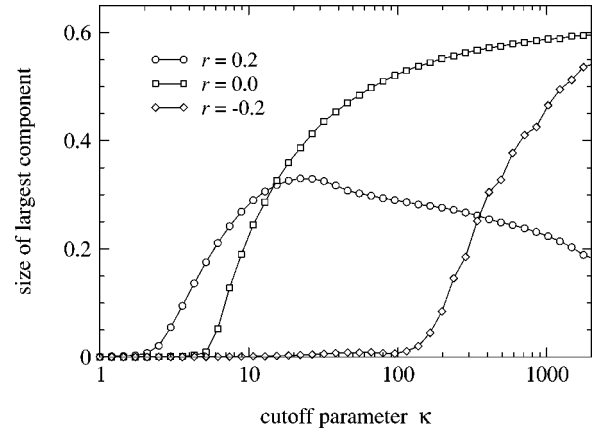

FIG. 3. Monte Carlo simulation results for the size of the giant component in graphs with the degree distribution Eq. (33), as a function of the cutoff parameter κ, with $\tau = 2.5$ and $\kappa' = 2$. The individual curves are for different values of the assortativity, as marked. Each data point is an average over 100 graphs of 100 000 vertices.

generally, we would like to be able to vary r freely, keeping the degree distribution fixed. We can do this by writing e_{jk} in the form

$$e_{jk} = q_j q_k + r \sigma_q^2 m_{jk}, \qquad (35)$$

where the symmetric matrix m_{jk} has all row and column sums zero and is normalized such that

$$\sum_{jk} jk m_{jk} = 1. \qquad (36)$$

For any choice of m_{jk} satisfying these constraints, Eq. (35) gives us a one-parameter family of networks parametrized by the assortativity coefficient r. We can for example choose

$$m_{jk} \sim q_j q_k - e_{jk}^{(d)} = (q_j - x_j)(q_k - x_k), \qquad (37)$$

for any correctly normalized x_k. Then Eq. (35) allows us to interpolate smoothly between $e_{jk}^{(d)}$ and $e_{jk}^{(a)}$ (and beyond) by simply varying the value of r. Note that whenever $r = 0$, we get a simple random graph without degree correlations, of the type discussed by Molloy and Reed [32] and others [66].

C. Simulation results

For e_{jk} chosen according to Eqs. (35) and (37), with q_k and x_k taking the same truncated power-law form as in Fig. 2, we have performed simulations for a variety of values of the parameters τ, κ, and κ'.

In Fig. 3 we show the size of the largest component in the graph as a function of κ, for three different values of the assortativity coefficient r. As the cutoff parameter κ increases, the mean degree of the graph increases also, so that the graph becomes more dense, ultimately passing the critical point at which a giant component develops. The figure reveals two findings of particular note.

(1) The position of the phase transition at which the giant-component appears moves to higher values of κ as the value

of r decreases. That is, the more assortative a network is, the lower the edge density at which a giant component first appears. This is intuitively reasonable. In assortatively mixed networks, the high-degree vertices tend to associate preferentially with one another, sticking together and forming what in the epidemiological literature is called a "core group." Within this core group the edge density is higher than it is in the graph as a whole, since the vertices in the group have higher-than-average degree. Thus one would expect to see a giant-component forming in the core group before it would form in a graph of the same average density but with no assortative mixing. Conversely, in graphs that are disassortatively mixed, the phase transition happens at a higher density than in neutrally mixed graphs.

(2) The size of the giant component in the limit of large κ is smaller for the assortatively mixed graph than for the neutral and disassortative ones. While this might seem at first to be at odds with the result that assortative graphs show a phase transition at lower density, it is really a reflection of the same underlying mechanism. Although the presence of a core group in an assortative graph allows it to percolate at a lower average density than other graphs, it also means that the density in other parts of the graph, outside the core group, is lower, and hence that the giant component is unlikely to extend into those regions. Thus the giant component is confined to the core of the network, and cannot grow as large as in a neutral or disassortative network.

A question of considerable interest in the study of networked systems is that of network resilience to the deletion of vertices. Suppose vertices are removed one by one from a network. How many must be removed before the giant component of the network is destroyed and network communication between distant vertices can no longer take place? Many networks, particularly those with highly skewed degree distributions, are found to be resilient to the random deletion of vertices but susceptible to the targeted deletion specifically of those vertices that have the highest degrees [8,9,11,67]. As we now show, these general results are modified by the pres-
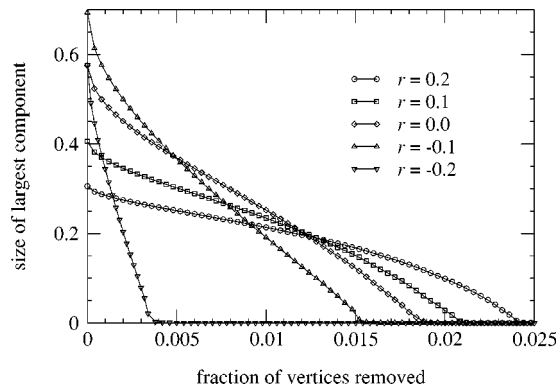
FIG. 4. The size of the largest component in a network as a fraction of system size as the highest-degree vertices are removed one by one. Each curve is for a single network of 10^7 vertices generated using the Monte Carlo method described in the text, with the degree distribution and the matrix e_{jk} chosen according to Eqs. (31) and (33), with $\tau=2.5$, $\kappa=500$, $\kappa'=5$, and r values as marked. Clearly the most assortatively mixed network ($r=0.2$) is considerably more robust against the removal of vertices than the most disassortative ($r=-0.2$).

ence of assortative mixing in the network.

In Fig. 4, we show the size of the largest component for five networks with different values of r as vertices are removed in decreasing order of their degree—i.e., highest-degree vertices first [68]. As the figure shows, although each of the networks has the same degree distribution, there is dramatic variation in the resilience of the networks with their assortativity. For the most assortative network, with $r=0.2$, it requires the removal of about ten times as many high-degree vertices to destroy the giant component as for the most disassortative one, with $r=-0.2$, even though the disassortative network starts out with a giant component about twice as big.

This finding fits naturally with our picture of an assortative network as dominated by a core group of interconnected high-degree vertices. Such a core group provides robustness to the network by concentrating all the obvious target vertices together in one portion of the graph. Removing these high-degree vertices is still one of the most effective ways to destroy network connectivity, but it is less effective because by removing them all from the same portion of the graph we fail to attack other portions. And if those other portions are themselves percolating, then a giant component will persist even as the highest-degree vertices vanish.

Conversely, the disassortatively mixed network is particularly susceptible to removal of high-degree vertices because those vertices are strewn far apart across the network, so that attacking them attacks all parts of the network at once.

One can also ask about the resilience of networks under random failure of their vertices (rather than targeted attack) [8,10,11]. Although we do not treat this case in detail here, it is reasonable to suppose that it is similar to the case of targeted attack. If assortative mixing makes networks more resilient against removal of their highest-degree vertices, then presumably they will also be resilient against removal of random ones; random vertex failure will do most damage

when it happens to hit high-degree vertices, but, as we have seen, this vulnerability is diminished by the concentration of the high-degree vertices in the core group. Some qualitative behaviors of the system may be unaffected by assortativity, however. For example, it is known that the fraction of vertices that must randomly fail to destroy the giant component in a network with a power-law degree distribution and uncorrelated degrees tends to unity as graph size becomes large, provided the exponent of the power-law satisfies $\tau<3$ [10]. Vazquez and Moreno [69] have recently shown that this result is not affected by the presence of assortative mixing by degree in the network, although disassortative mixing can make a difference.

D. Discussion

The results found here could have applications in a variety of areas. Consider for example, the spread of diseases on networks, which has been the subject of much attention in the recent networks literature [70–77]. The largest component of the contact network over which a disease spreads represents the largest possible disease outbreak on that network, and a network with no giant component cannot show an epidemic (system-wide) outbreak. Thus our finding, that a giant component forms more easily in a network that is assortatively mixed by degree, suggests that in such networks epidemic outbreaks would become possible at a lower edge density than in the corresponding disassortative network. In the language of epidemiology, the core group of an assortatively mixed network forms a "reservoir," which can sustain an outbreak of the disease even when the density of the network as a whole is too low to do so. On the other hand, the smaller asymptotic size of the giant component in an assortatively mixed network seems to imply that, when they occur, epidemics in such networks would be restricted to a smaller segment of the population than in a similar disassortative network—the outbreak is confined mostly to the core group and does not spread to the population as a whole. Thus from the epidemiological point of view there are both good and bad sides to the phenomenon of assortativity.

One could test these predictions explicitly by studying epidemic models such as SIR or SIRS models [78,79] on assortatively mixed model networks of the type introduced here. Some studies of this kind have already been carried out—see, for example, Refs. [44] and [80]—although the particular conjectures put forward here have not been conclusively verified.

Our findings on network resilience also have some practical implications. In the context of epidemiology, for instance, removal of vertices from the network might correspond to immunization of individuals to prevent the spread of disease. Assuming that the goal of a vaccination program is to destroy network connectivity so that the disease in question cannot spread, our findings suggest that even targeted vaccination strategies would be less effective in assortative networks than in disassortative or neutral ones because of the resilience of the network to this type of attack. In other contexts, however, resilience is a good thing. For example we would like to make networks such as the internet and other communication or distribution networks resilient against at-

tacks on their vertices. In this context assortative mixing would be beneficial.

Unfortunately, when we look at Table II, we find a discouraging picture. As we pointed out in Sec. III A, almost all the social networks we have looked at are significantly assortative, meaning that they would be robust to vertex removal. But these are the very networks by which disease spreads, those that we would like to be able to attack using vaccination strategies. Even the email network, which is relevant to the spread of computer viruses [50], is assortative and hence resilient. On the other hand, the technological networks like the Internet, which we would like to be able to protect, are disassortative, and hence particularly vulnerable to targeted attack.

IV. CONCLUSIONS

In this paper, we have studied the phenomenon of assortative mixing in networks, which is the tendency for vertices in networks to connect preferentially to other vertices that are like them in some way. This preference may take a number of forms. Mixing may follow discrete or enumerative characteristics. In the social networks that have been the main focus of this paper, connections between people may be assortative by language, for example, or by race—people may prefer to associate with others who speak the same language as they do or are of the same race. Mixing can also be dictated by scalar characteristics such as age or income. A special case of mixing by a scalar characteristic is mixing according to vertex degree, which has been shown previously to be present in a variety of networks, including nonsocial ones such as the Internet and protein interaction networks. Mixing can also be disassortative, meaning that vertices in the network preferentially form connections to others that are unlike them.

We have proposed some simple measures for these types of mixing, which we call assortativity coefficients. These measures are positive or negative for assortative or disassortative mixing, respectively, and zero for neutrally mixed networks. Applying our measures to a broad selection of network data drawn from various real-world situations we have shown that the phenomenon of assortative mixing is indeed widespread, with only a few of the networks studied showing no statistically significant biases in their mixing patterns. In the case of mixing by vertex degree, a remarkable pattern emerges. Almost all the social networks studied show positive assortativity coefficients while all other types of networks, including technological and biological networks, show negative coefficients, i.e., disassortative mixing. Only three networks that showed no significant trend either way failed to follow this rule. We have offered some conjectures about the origin of this striking regularity, but we believe it unlikely that any single mechanism can explain the mixing patterns of all of these disparate networks.

We have also proposed a number of models of assortatively mixed networks, for mixing both by discrete and by scalar characteristics. For each of the mixing types considered it is possible to create random graph models for which one can calculate exactly by generating function methods certain average properties of network ensembles. We have also described Monte Carlo methods for generating random graphs drawn from each of the classes discussed with specified values of the mixing parameters.

For the case of mixing by vertex degree we have performed extensive simulations. Two results of particular interest emerge from these studies. First, we find that networks that are assortatively mixed by degree percolate more easily that their disassortative counterparts. That is, a giant component of connected vertices forms in the network at lower edge density than in another network with the same degree distribution but zero or negative assortativity. This result may imply, for instance, that assortatively mixed social networks would support epidemic disease outbreaks more easily than disassortatively mixed ones, which would be a disheartening conclusion, given our finding that most social networks appear to be assortative.

Second, we find that assortatively mixed networks are more robust to the deletion of their vertices than disassortatively mixed or neutral networks. We have studied in particular the case of the targeted deletion of the highest-degree vertices, which has been suggested as a possible vaccination strategy for breaking up networks of disease-causing contacts, but it is reasonable to suppose that the same result will extend also to the random deletion of vertices. This result too leads to a rather gloomy conclusion: targeted vaccination strategies may be less effective than we would hope in preventing disease because of the assortative and hence resilient nature of social networks, while on the other hand networks that we would hope to protect against vertex removal, communication networks like the internet, for instance, will be particularly susceptible because of their disassortative nature.

APPENDIX A: THE ASSORTATIVITY MEASURE OF GUPTA, ANDERSON, AND MAY

Gupta, Anderson, and May [26] have defined a measure of assortative mixing by discrete types different from the one that we have introduced in Sec. II A. In our notation their measure is

$$Q = \frac{\sum_i a_i^{-1}(e_{ii} - a_i b_i)}{n-1} = \frac{\sum_i e_{ii}/a_i - 1}{n-1}, \quad \text{(A1)}$$

where as before n is the number of vertex types, and we have made use of $\sum_i b_i = 1$. Like our measure, this measure is 0 for a neutrally mixed network and 1 for a perfectly assortative

network. In general, however, the values of the two measures are quite different. Here we give a simple example to illustrate the difference between the two.

Consider a network with three types of vertices. There are 100 vertices of type 1, 100 of type 2, and 2 of type 3. The vertices of types 1 and 2 mix indiscriminately with one other—connections from 1 to 2 are as likely as from 1 to 1, and so forth. The vertices of type 3 however, associate only among themselves and not with types 1 and 2 at all. This is reflected in the matrix **e** for the 202 vertices thus

$$
\mathbf{e} = \frac{1}{202} \begin{pmatrix} 50 & 50 & 0 \\ 50 & 50 & 0 \\ 0 & 0 & 2 \end{pmatrix}.
\tag{A2}
$$

Clearly most of this network—99% of it, in fact—is mixing randomly, and hence we would expect the assortativity coefficient to be close to zero. The value of r for the matrix above

reflects this; we find $r = 0.029$. The measure of Gupta, Anderson, and May [26], however, takes a value $Q = 0.50$. This appears to indicate that the network has very strong assortative mixing, when in fact it does not. The reason for this is that the measure of Gupta, Anderson, and May, rather than giving each vertex in the network equal weight, weights each *type* of vertex equally, so that vertices that belong to large groups get less weight in the calculation than those in small groups. In the present case, where one group is very small, the vertices in that group are weighted very heavily, and since those vertices mix perfectly assortatively, the value of Q is, as a result, high. If we remove these vertices from the network, the value of Gupta, Anderson, and May's Q coefficient jumps to zero. Thus the two vertices in the third group have a disproportionately large effect on the value of Q.

The solution to this problem is to give each vertex equal weight in the calculation, which is precisely what our measure r does.

---

[1] S.H. Strogatz, Nature (London) **410**, 268 (2001).

[2] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[3] S.N. Dorogovtsev and J.F.F. Mendes, Adv. Phys. **51**, 1079 (2002).

[4] J. Travers and S. Milgram, Sociometry **32**, 425 (1969).

[5] D.J. Watts and S.H. Strogatz, Nature (London) **393**, 440 (1998).

[6] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[7] L.A.N. Amaral, A. Scala, M. Barthélémy, and H.E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11 149 (2000).

[8] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).

[9] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Comput. Netw. **33**, 309 (2000).

[10] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).

[11] D.S. Callaway, M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).

[12] P. Holme, B.J. Kim, C.N. Yoon, and S.K. Han, Phys. Rev. E **65**, 056109 (2002).

[13] J. M. Kleinberg, in *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing* (Association of Computing Machinery, New York, 2000), pp. 163–170.

[14] L.A. Adamic, R.M. Lukose, A.R. Puniyani, and B.A. Huberman, Phys. Rev. E **64**, 046135 (2001).

[15] D.J. Watts, P.S. Dodds, and M.E.J. Newman, Science **296**, 1302 (2002).

[16] M. Girvan and M.E.J. Newman, Proc. Natl. Acad. Sci. U.S.A. **99**, 8271 (2002).

[17] P. Holme, M. Huss, and H. Jeong, e-print cond-mat/0206292.

[18] R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt, and A. Arenas, e-print cond-mat/0211498.

[19] R. Monasson, Eur. Phys. J. B **12**, 555 (1999).

[20] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. E **64**, 051903 (2001).

[21] I.J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, Phys. Rev. E **64**, 026704 (2001).

[22] M.E.J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).

[23] We find it convenient even on a undirected graph to consider the ends of the edges to be distinguishable—each edge has a unique A-end and B-end, which are marked in some way. We can think of one of the ends as having a dot or other identifying feature on it. This makes the counting of edges simpler: the matrix element e_{ij} is defined as the probability that a randomly chosen edge is connected to a vertex of type i at its A end and type j at its B end. Thus every edge, whether it joins unlike vertices or like ones, appears only once in the matrix—no edge appears both above and below the diagonal. It is possible to construct a theory in which the ends of undirected edges are indistinguishable, but in this case each edge that joins unlike vertices appears twice in the matrix, both above and below the diagonal, and edges joining like vertices appear only once. This necessitates the introduction of an extra factor of 2 into the off-diagonal terms. This approach is adopted for example in Ref. [43].

[24] J.A. Catania, T.J. Coates, S. Kegels, and M.T. Fullilove, Am. J. Public Health **82**, 284 (1992).

[25] M. Morris, in *Epidemic Models: Their Structure and Relation to Data*, edited by D. Mollison (Cambridge University Press, Cambridge, 1995), pp. 302–322.

[26] S. Gupta, R.M. Anderson, and R.M. May, AIDS **3**, 807 (1989).

[27] B. Efron, SIAM Rev. **21**, 460 (1979).

[28] J. L. Fleiss, *Statistical Methods for Rates and Proportions* (Wiley, New York, 1981).

[29] J.L. Fleiss, J. Cohen, and B.S. Everitt, AAS Photo-Bull. **72**, 323 (1969).

[30] E.A. Bender and E.R. Canfield, J. Comb. Theory, Ser. A **24**, 296 (1978).

[31] T. Łuczak, in *Proceedings of the Symposium on Random Graphs, Poznań 1989*, edited by A. M. Frieze and T. Łuczak (Wiley, New York, 1992), pp. 165–182.

[32] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).

[33] M. Molloy and B. Reed, Combinatorics, Probab. Comput. **7**, 295 (1998).

[34] W. Aiello, F. Chung, and L. Lu, in *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing* (Association of Computing Machinery, New York, 2000), pp. 171–180.

[35] M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Phys. Rev. E **64**, 026118 (2001).

[36] M.E.J. Newman, D.J. Watts, and S.H. Strogatz, Proc. Natl. Acad. Sci. U.S.A. **99**, 2566 (2002).

[37] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Oxford University Press, Oxford, 1999).

[38] *National Survey of Family Growth, Cycle V, 1995* (U.S. Department of Health and Human Services, National Center for Health Statistics, Hyattsville, MD, 1997).

[39] Perhaps it is stretching a point a little to regard links between marriage partners as forming a network, since presumably most people have only one marriage at a time. However, if we view the ages of marriage partners as a guide to the ages of sexual partners in general, then the resulting assortative mixing also describes networks of such more general partnerships, which are certainly very real [25,26,49,64].

[40] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).

[41] S. Maslov and K. Sneppen, e-print cond-mat/0205379.

[42] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, e-print cond-mat/0206467.

[43] A. Vazquez and M. Weigt, e-print cond-mat/0207035.

[44] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. E. (to be published), e-print cond-mat/0208163.

[45] One can also calculate a value for r by simply ignoring the directed nature of the edges in a directed network. This approach, which we adopted in Ref. [22], will in general give a different figure from that given by Eq. (25). While Eq. (25) will normally give a more meaningful result for a directed network, there may be cases in which ignoring direction is the correct thing to do. For example, in a food web one might only be interested in which species have tropic relations with with others, and not in which direction that relation lies in terms of energy or carbon flow.

[46] M.E.J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001).

[47] J.W. Grossman and P.D.F. Ion, Congr. Numer. **108**, 129 (1995).

[48] G. F. Davis, M. Yoo, and W. E. Baker (unpublished).

[49] P. S. Bearman, J. Moody, and K. Stovel (unpublished).

[50] M.E.J. Newman, S. Forrest, and J. Balthrop, Phys. Rev. E **66**, 035101 (2002).

[51] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. J. Shenker, and W. Willinger, in *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies* (IEEE Comput. Soc., Los Alanitos, CA, 2002).

[52] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).

[53] H. Jeong, S. Mason, A.-L. Barabási, and Z.N. Oltvai, Nature (London) **411**, 41 (2001).

[54] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.-L. Barabási, Nature (London) **407**, 651 (2000).

[55] J.G. White, E. Southgate, J.N. Thompson, and S. Brenner, Philos. Trans. R. Soc. London, Ser. A **314**, 1 (1986).

[56] M. Huxham, S. Beaney, and D. Raffaelli, Oikos **76**, 284 (1996).

[57] N.D. Martinez, Ecol. Monogr. **61**, 367 (1991).

[58] D. Strauss, SIAM Rev. **28**, 513 (1986).

[59] T.A.B. Snijders, J. Soc. Struct. **2** (2002).

[60] Strictly these probabilities are only correct in a "canonical ensemble" of graphs in which the degree distribution is fixed rather than the degree sequence. This ensemble and the fixed-degree-sequence one studied here, however, become equivalent in the limit of large graph size; the error introduced here by substituting one for the other is of the order of N^{-1} and is small compared with other sources of error in our simulations.

[61] D.J. de S. Price, Science **149**, 510 (1965).

[62] S. Redner, Eur. Phys. J. B **4**, 131 (1998).

[63] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999).

[64] F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley, and Y. Åberg, Nature (London) **411**, 907 (2001).

[65] H. Ebel, L.-I. Mielsch, and S. Bornholdt, Phys. Rev. E **66**, 035103 (2002).

[66] Not all graphs with $r=0$ are without degree correlations. A measurement of $r=0$ implies only that the mean degree correlation is zero when averaged over all degrees. The grown graph model of Barabási and Albert [6] provides an example of a network that possesses degree correlations although it has $r=0$ [22].

[67] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001).

[68] The degree is not recalculated after each removal. Removal is in the order of vertices' starting degree in the network before any deletion has taken place.

[69] A. Vazquez and Y. Moreno, Phys. Rev. E **67**, 015101 (2003).

[70] F. Ball, D. Mollison, and G. Scalia-Tomba, Ann. Appl. Probab. **7**, 46 (1997).

[71] H. Andersson, Math. Scientist **24**, 128 (1999).

[72] C. Moore and M.E.J. Newman, Phys. Rev. E **61**, 5678 (2000).

[73] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).

[74] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **63**, 066117 (2001).

[75] R.M. May and A.L. Lloyd, Phys. Rev. E **64**, 066112 (2001).

[76] L.M. Sander, C.P. Warren, I. Sokolov, C. Simon, and J. Koopman, Math. Biosci. **180**, 293 (2002).

[77] M.E.J. Newman, Phys. Rev. E **66**, 016128 (2002).

[78] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and its Applications* (Hafner Press, New York, 1975).

[79] H.W. Hethcote, SIAM Rev. **42**, 599 (2000).

[80] A. Vazquez and Y. Moreno, e-print cond-mat/0210362.