

Spectral graph analysis of modularity and assortativity

P. Van Mieghem,^{1,*} X. Ge,² P. Schumm,³ S. Trajanovski,¹ and H. Wang¹

¹*Faculty of Electrical Engineering, Mathematics, and Computer Science, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands*

²*College of Information Science and Engineering, Northeastern University, 110004 Shenyang, China*

³*Department of Electrical and Computer Engineering, Kansas State University, Manhattan, Kansas 66506, USA*

(Received 31 August 2010; revised manuscript received 15 October 2010; published 16 November 2010)

Expressions and bounds for Newman's modularity are presented. These results reveal conditions for or properties of the maximum modularity of a network. The influence of the spectrum of the modularity matrix on the maximum modularity is discussed. The second part of the paper investigates how the maximum modularity, the number of clusters, and the hop count of the shortest paths vary when the assortativity of the graph is changed via degree-preserving rewiring. Via simulations, we show that the maximum modularity increases, the number of clusters decreases, and the average hop count and the effective graph resistance increase with increasing assortativity.

DOI: [10.1103/PhysRevE.82.056113](https://doi.org/10.1103/PhysRevE.82.056113)

PACS number(s): 89.75.Fb, 02.10.-v

I. INTRODUCTION

Graph communities reveal important structural features of the topology. The communities can be observed as sections of the graph topology that exhibit relatively higher levels of connections within the regions and lower connectivity between the regions. Such a structure plays a significant role in both the sorting of nodes and the evolution of processes on graphs, including slowing spreading epidemics and containing cascading failures [1–3]. Although there is no definition of community that is accepted for all graph-based systems, a metric known as modularity has led to a surge of research in community structure discovery and analysis [4]. Fortunato [5] provides a sound summary of this collection of works. A significant fraction of this research focuses on ways to assign nodes to communities in efforts to maximize the modularity metric [6]. Other pursuits have begun characterizing the modularity metric, investigating the counterintuitive non-trivial expected values for random graph models and lattices [7] and the upper bound and partitioning resolution [8–13]. Characterizing modularity is also done by examining how it relates to other significant graph metrics. Introduced only shortly before modularity, assortativity is a correlation of the similarities of nodes sharing a link [14,15]. Newman suggested that one driving factor in the formation of communities was the preference of nodes to connect to other nodes that possessed similar characteristics to their own. This has been observed in some social networks where the similarities are race related [16]. Within topological analysis, assortativity is most commonly used with the node degrees or node strengths. Previous work has shown relations between the assortativity of a graph and the characteristic path length, the fraction of nodes in the giant component, the clustering coefficient, the robustness, the spectra of the adjacency matrix, and the modularity [17–20].

This paper presents an analysis of the relationships among the modularity, the assortativity, the largest eigenvalues of

the adjacency and modularity matrices, the number of clusters, the hop count or characteristic path length, and the effective graph resistance. Section II summarizes mathematical formulations of the modularity and it provides bounds. Section III attempts to understand the effect of increasing and decreasing the assortativity of a graph via degree-preserving rewiring. We show that, in most complex networks, the maximum modularity (and the largest eigenvalue of the modularity matrix) increases with assortativity.

II. MODULARITY

The modularity m , proposed by Newman [4], is a measure of the quality of a particular division of the network, which is defined in [21] as¹

$$m = \frac{1}{2L} \sum_{i=1}^N \sum_{j=1}^N \left(a_{ij} - \frac{d_i d_j}{2L} \right) 1_{\{i \text{ and } j \text{ belong to the same community}\}}, \quad (1)$$

where a_{ij} is the matrix element of the adjacency matrix A of the graph with N nodes and L links. The modularity is proportional to the number of links falling within clusters or groups minus the expected number in an equivalent network with links placed at random. Thus, if the number of links within a group is not better than random, the modularity is zero. A modularity approaching 1 reflects networks with strong community structure: a dense intragroup and a sparse intergroup connection pattern. If links are placed at random, then the expected number of links between node i and node j equals $d_i d_j / 2L$, where d_j is the degree of node j .

A. Expressions and bounds for the modularity

The general definition (1) is first rewritten as follows. We transform the nodal representation to a counting over links $l=i \sim j$ such that

*p.f.a.vanmieghem@tudelft.nl

¹We follow the notation of our book [22].

$$\sum_{i=1}^N \sum_{j=1}^N a_{ij} 1_{\{i \text{ and } j \text{ belong to the same cluster}\}} = 2 \sum_{k=1}^c L_k,$$

where L_k is the number of links of cluster C_k , and the factor 2 arises from the fact that all links are counted twice due the symmetry $A=A^T$ of the adjacency matrix. If we denote by L_{inter} the number of intercommunity links, i.e., the number of links that are cut by partitioning the network into c communities or clusters, then

$$L = \sum_{k=1}^c L_k + L_{\text{inter}}.$$

Similarly,

$$\begin{aligned} & \sum_{i=1}^N \sum_{j=1}^N d_i d_j 1_{\{i \text{ and } j \text{ belong to the same cluster}\}} \\ &= \sum_{k=1}^c \left(\sum_{i \in C_k} d_i \right) \left(\sum_{j \in C_k} d_j \right) = \sum_{k=1}^c D_{C_k}^2, \end{aligned}$$

where

$$D_{C_k} = \sum_{i \in C_k} d_i$$

is the sum of the degrees of all nodes that belong to cluster C_k . Clearly, $D_{C_k} \geq 2L_k$, because some nodes in cluster C_k may possess links connected to nodes in other clusters. The basic law of the degree then shows that $\sum_{k=1}^c D_{C_k} = 2L$. Substituting these expressions in definition (1) leads to an alternative expression² for the modularity:

$$m = \sum_{k=1}^c \left[\frac{L_k}{L} - \left(\frac{D_{C_k}}{2L} \right)^2 \right]. \quad (2)$$

The last fraction can also be written as

$$\frac{D_{C_k}}{2L} = \frac{n_k E[D_k]}{N E[D]},$$

where n_k and D_k are the number of nodes and the degree of a random node in cluster C_k , respectively. Thus, the modularity is the sum over all clusters of the fraction of links per cluster minus the square of the fraction of nodes per cluster multiplied by the ratio of the average degree in a cluster over the average degree in the network.

Subject to the basic law of the degree, $\sum_{k=1}^c D_{C_k} = 2L$, the sum $\sum_{k=1}^c D_{C_k}^2$ is maximized when $D_{C_k} = \frac{2L}{c}$ for all $1 \leq k \leq c$. Indeed, the corresponding Lagrangian,

$$\mathcal{L} = \sum_{k=1}^c D_{C_k}^2 + \xi \left(\sum_{k=1}^c D_{C_k} - 2L \right),$$

where ξ is a Lagrange multiplier, supplies the set of equations for the optimal solution, $\partial \mathcal{L} / \partial D_{C_j} = 2D_{C_j} + \xi = 0$ for $1 \leq j \leq c$ and $\frac{\partial \mathcal{L}}{\partial \xi} = \sum_{k=1}^c D_{C_k} - 2L = 0$, which is satisfied for

$\xi = -\frac{4L}{c}$ and $D_{C_j} = \frac{2L}{c}$ for all $1 \leq j \leq c$. Hence, $\sum_{k=1}^c D_{C_k}^2 \leq (2L)^2/c$. The modularity in Eq. (2) is minimized, for $c > 1$, if $L_k = 0$ for $1 \leq k \leq c$, and $\sum_{k=1}^c D_{C_k}^2$ is maximized such that $m \geq -\frac{1}{c}$. In conclusion, the modularity of any graph is never smaller than $-\frac{1}{2}$, and this minimum is obtained for the complete bipartite graph. This result was earlier stated in [23], but with a different proof.

Invoking the Cauchy identity (see [22]) and $\sum_{k=1}^c D_{C_k} = 2L$,

$$\sum_{k=1}^c D_{C_k}^2 = \frac{(2L)^2}{c} + \frac{1}{c} \sum_{j=2}^c \sum_{k=1}^{j-1} (D_{C_j} - D_{C_k})^2,$$

results in yet another expression for the modularity:

$$m = 1 - \frac{L_{\text{inter}}}{L} - \frac{1}{c} - \frac{1}{c} \sum_{j=2}^c \sum_{k=1}^{j-1} \left(\frac{D_{C_j} - D_{C_k}}{2L} \right)^2. \quad (3)$$

Since the double sum is always positive, Eq. (3) provides us with an upper bound for the modularity,

$$m \leq 1 - \frac{1}{c} - \frac{L_{\text{inter}}}{L}. \quad (4)$$

The upper bound (4) is only attained if the degree sum of all clusters is the same. In passing, we mention that Eq. (3) rigorously proves the upper bound derived by Fortunato and Barthélemy [10] based on a cyclic chain of same subgraphs, for which, indeed, $D_{C_j} = D_{C_k}$ for each pair (j, k) . In addition, the upper bound (4) shows that $m \leq 1$ and that a modularity of 1 is only reached asymptotically, when the number of clusters $c \rightarrow \infty$ and $L_{\text{inter}} = o(L)$, implying that the fraction of intercommunity links over the total number of links L is vanishingly small for large graphs ($N \rightarrow \infty$ and $L \rightarrow \infty$).

Let $D_{\Delta C} = \max_{\{C_j, C_k\}} |D_{C_j} - D_{C_k}|$, then a lower bound of the modularity, deduced from Eq. (3), is

$$m \geq 1 - \frac{L_{\text{inter}}}{L} - \frac{1}{c} - \frac{(c-1)}{2} \left(\frac{D_{\Delta C}}{2L} \right)^2. \quad (5)$$

Only if $D_{\Delta C} = 0$ the lower bound (5) equals the upper bound (4) and the equality sign can occur. Excluding the case that $D_{\Delta C} = 0$, then not all D_{C_j} 's are equal, and we may assume an ordering $D_{C_1} \geq D_{C_2} \geq \dots \geq D_{C_c}$, with at least one strict inequality. We demonstrate that, for $c > 2$, not all differences $D_{C_j} - D_{C_k} = D_{\Delta C} > 0$ for any pair (j, k) . For, assume the contrary, so that $D_{C_1} - D_{C_2} = D_{C_2} - D_{C_3} = D_{C_1} - D_{C_3} = D_{\Delta C} > 0$, then $D_{\Delta C} = D_{C_1} - D_{C_3} = (D_{C_1} - D_{C_2}) + (D_{C_2} - D_{C_3}) = 2D_{\Delta C}$, which cannot hold for $D_{\Delta C} > 0$. Hence, if $D_{\Delta C} > 0$, the inequality in Eq. (5) is strict; alternatively, the lower bound (5) is not attainable in that case.

In order for a network to have a modular structure, the modularity must be positive. The requirement that the lower bound (5) is non-negative supplies us with an upper bound for the maximum difference $D_{\Delta C}$ in the nodal degree sum between two clusters in a ‘‘modular’’ graph,

²Newman [33] presented still another expression for the modularity.

$$D_{\Delta C} \leq 2L \sqrt{\frac{2}{c-1} \left(1 - \frac{L_{\text{inter}}}{L} - \frac{1}{c}\right)}. \quad (6)$$

For $c > 1$, Eq. (6) demonstrates that $D_{\Delta C} < 2L$. Ignoring the integer nature of c , the lower bound (5) is maximized with respect to the number of clusters c when

$$c^* = \frac{2\sqrt{2}L}{D_{\Delta C}} > \sqrt{2}, \quad (7)$$

resulting in

$$m \geq 1 - \frac{L_{\text{inter}}}{L} - \sqrt{2} \left(\frac{D_{\Delta C}}{2L} \right) + \frac{1}{2} \left(\frac{D_{\Delta C}}{2L} \right)^2.$$

The right-hand side in this lower bound is positive provided that $1 > D_{\Delta C}/2L > \sqrt{2}(1 - \sqrt{L_{\text{inter}}/L})$. When this lower bound for $D_{\Delta C}/2L$ is satisfied, the modularity m is certainly positive, implying that the graph exhibits a modular structure.

Another presentation for the modularity applies the identity

$$\left(\sum_{j=1}^n x_j \right)^2 = \sum_{j=1}^n \sum_{k=1}^n x_j x_k = \sum_{j=1}^n x_j^2 + 2 \sum_{j=2}^n \sum_{k=1}^{j-1} x_j x_k \quad (8)$$

to $x_j = D_{C_j}$,

$$\sum_{k=1}^c D_{C_k}^2 = (2L)^2 - 2 \sum_{j=2}^c \sum_{k=1}^{j-1} D_{C_j} D_{C_k},$$

such that Eq. (2) is rewritten as

$$m = \frac{2}{(2L)^2} \sum_{j=2}^c \sum_{k=1}^{j-1} D_{C_j} D_{C_k} - \frac{L_{\text{inter}}}{L}. \quad (9)$$

Using the basic law of the degree, $\sum_{k=1}^c D_{C_k} = 2L$, the first term in Eq. (9) is maximized, as follows from a similar Lagrangian argument as before, when the degree is distributed uniformly across the communities as $D_{C_k} = 2L/c$, resulting in

$$\begin{aligned} m &\leq \frac{1}{2L^2} \sum_{j=2}^c (j-1) \left(\frac{2L}{c} \right)^2 - \frac{L_{\text{inter}}}{L} \\ &\leq \frac{1}{2L^2} \left(\frac{2L}{c} \right)^2 \left(\frac{c^2 - c}{2} \right) - \frac{L_{\text{inter}}}{L} \leq 1 - \frac{1}{c} - \frac{L_{\text{inter}}}{L}, \end{aligned}$$

which is again the upper bound (4) and, hence, agrees with the degree balancing of Eq. (3). Equations (3) and (9) present the maximization of the modularity from dual perspectives, yet both point to a common solution of degree balancing and minimizing L_{inter} .

Finally, we present a probabilistic setting for the modularity m by defining the random variable D_C as the sum of the degrees in an arbitrary cluster. The average is $E[D_C] = \frac{1}{c} \sum_{k=1}^c D_{C_k} = \frac{2L}{c}$ and, comparing with the estimate c^* in Eq. (7), it always holds that $c = 2L/E[D_C]$. However, estimate (7) suggests that the extreme difference $D_{\Delta C}$ is not that far way from the mean, roughly by a factor of $\sqrt{2}$. Further, with

$$\frac{1}{c} \sum_{k=1}^c D_{C_k}^2 = \text{Var}[D_C] + (E[D_C])^2,$$

expression (2) for the modularity becomes

$$m = 1 - \frac{L_{\text{inter}}}{L} - \frac{1}{c} - \frac{c}{(2L)^2} \text{Var}[D_C], \quad (10)$$

which, again, leads to the upper bound (4) when the variance is zero, i.e., when all clusters have an equal degree sum. Incidentally, comparing Eqs. (3) and (10), we find that

$$\text{Var}[D_C] = \sum_{j=2}^c \sum_{k=1}^{j-1} \left(\frac{D_{C_j} - D_{C_k}}{c} \right)^2,$$

and this is a general result that holds for any random variable in a specific graph [22].

B. Spectral form for the modularity

The $N \times c$ community matrix S , defined as

$$S_{ik} = \begin{cases} 1 & \text{if node } i \text{ belongs to community } k \\ 0 & \text{otherwise,} \end{cases}$$

can be used to rephrase condition in (1) as

$$\mathbf{1}_{\{i \text{ and } j \text{ belong to the same community}\}} = \sum_{k=1}^c S_{ik} S_{jk},$$

leading to the matrix representation of the modularity:

$$m = \frac{1}{2L} \sum_{k=1}^c \sum_{i=1}^N \sum_{j=1}^N S_{ik} m_{ij} S_{jk} = \frac{\text{tr}(S^T M S)}{2L}, \quad (11)$$

where

$$M = A - \frac{1}{2L} d \cdot d^T \quad (12)$$

is the modularity matrix and d is the degree vector. We define the community vector s_k , which equals the k th column of the community matrix S and which specifies the k th cluster: all components of s_k , corresponding to nodes belonging to cluster C_k , are equal to 1; otherwise, they are zero.

Using the eigenvalue decomposition of the symmetric modularity matrix $M = W \text{diag}[\lambda_j(M)] W^T$, where W is the orthogonal $N \times N$ matrix with the j th eigenvector w_j belonging to $\lambda_j(M)$ in column j , the general spectral expression for the modularity m follows from Eq. (11) as

$$\begin{aligned} m &= \frac{\text{tr}\{(W^T S)^T \text{diag}[\lambda_j(M)] W^T S\}}{2L} \\ &= \frac{1}{2L} \sum_{j=1}^N \left(\sum_{k=1}^c (w_j^T s_k)^2 \right) \lambda_j(M), \end{aligned} \quad (13)$$

because $(W^T S)_{jk} = \sum_{l=1}^N W_{lj} S_{lk} = w_j^T s_k$. In particular, the scalar product $w_j^T s_k = \sum_{q \in C_k} (w_j)_q$ is the sum of those eigenvector components of w_j that belong to cluster C_k . If we write the community vector as a linear combination of the eigenvec-

tors of M , $s_k = \sum_{j=1}^N \beta_{kj} w_j$, then the orthogonality of eigenvectors indicates that the coefficients equal $\beta_{kj} = w_j^T s_k$. Moreover, the vectors s_1, s_2, \dots, s_c are orthogonal vectors; and, by definition, $\sum_{k=1}^c s_k = u$. Since u is an eigenvector of M belonging to the zero eigenvalue as follows from the definition of the modularity matrix (12) because

$$M \cdot u = Au - \frac{1}{2L} d \cdot d^T u = d - d = 0,$$

and $Au = d$ and $d^T u = 2L$, we observe that

$$\sum_{k=1}^c w_j^T s_k = 0,$$

provided the eigenvector $w_j \neq u$. Using the Cauchy identity

$$\begin{aligned} c \sum_{k=1}^c (w_j^T s_k)^2 - \left(\sum_{k=1}^c w_j^T s_k \right)^2 &= \frac{1}{2} \sum_{m=1}^c \sum_{k=1}^c [w_j^T (s_m - s_k)]^2 \\ &= \sum_{m=2}^c \sum_{k=1}^{m-1} [w_j^T (s_m - s_k)]^2, \end{aligned}$$

we find that

$$m = \frac{1}{2Lc} \sum_{j=1}^N \left(\sum_{m=2}^c \sum_{k=1}^{m-1} [w_j^T (s_m - s_k)]^2 \right) \lambda_j(M). \quad (14)$$

For $c=2$ and $y = s_1 - s_2$, which is a vector with component $y_j = 1$ if node j belongs to cluster C_1 and $y_j = -1$ if node j belongs to cluster C_2 , the general relation (14) reduces to

$$m_2 = \frac{1}{4L} \sum_{j=1}^N \beta_j^2 \lambda_j(M), \quad (15)$$

where $y = \sum_{j=1}^N \beta_j w_j$, with $\beta_j = y^T w_j$. Expression (15) was Newman's starting point in [21] for his iterated bisection method.

Since $WW^T = I$, we have that $\text{tr}[(W^T S)^T W^T S] = \text{tr}(S^T S) = N$ (see [22]), such that

$$\sum_{j=1}^N \sum_{k=1}^c (w_j^T s_k)^2 = N. \quad (16)$$

In the bicluster case where $c=2$, we see that $y^T y = N$ such that $\sum_{j=1}^N \beta_j^2 = N$. Let $w_q = \frac{u}{\sqrt{N}}$ denote the eigenvector of M belonging to the eigenvalue $\lambda_q(M) = 0$, then

$$\sum_{k=1}^c (w_q^T s_k)^2 = \frac{1}{N} \sum_{k=1}^c (u^T s_k)^2 = \frac{1}{N} \sum_{k=1}^c n_k^2,$$

where n_k is the number of nodes in cluster C_k . By applying the inequality

$$\min_{1 \leq k \leq n} \frac{a_k}{q_k} \leq \frac{a_1 + a_2 + \dots + a_n}{q_1 + q_2 + \dots + q_n} \leq \max_{1 \leq k \leq n} \frac{a_k}{q_k}, \quad (17)$$

where q_1, q_2, \dots, q_n are positive real numbers and a_1, a_2, \dots, a_n are real numbers,

$$\begin{aligned} \frac{\sum_{j=1; j \neq q}^N \left(\sum_{k=1}^c (w_j^T s_k)^2 \right) \lambda_j(M)}{\sum_{j=1; j \neq q}^N \sum_{k=1}^c (w_j^T s_k)^2} &\leq \max_{1 \leq j \leq N} \frac{\left(\sum_{k=1}^c (w_j^T s_k)^2 \right) \lambda_j(M)}{\sum_{k=1}^c (w_j^T s_k)^2} \\ &= \lambda_1(M), \end{aligned}$$

from which we find, with $E[D] = \frac{2L}{N}$, a spectral upper bound for the modularity:

$$m \leq \frac{\lambda_1(M)}{E[D]} \left(1 - \frac{1}{N^2} \sum_{k=1}^c n_k^2 \right).$$

This bound can also be written as

$$m \leq \frac{\lambda_1(M)}{E[D]} \left(1 - \frac{1}{c} - \frac{c}{N^2} \text{Var}[n_C] \right),$$

where n_C is the number of nodes in an arbitrary cluster, because $E[n_C] = \frac{1}{c} \sum_{k=1}^c n_k = \frac{N}{c}$. Since $\text{Var}[n_C] \geq 0$, we arrive at the upper bound

$$m \leq \frac{\lambda_1(M)}{E[D]} \left(1 - \frac{1}{c} \right). \quad (18)$$

We observe that Eq. (18) may lead to a sharper upper bound than Eq. (4) if $\lambda_1(M) < E[D]$ (see, e.g., Fig. 3 below).

We have shown in [22] that the eigenvalues of the modularity matrix $M = A - \frac{1}{2L} d \cdot d^T$ are interlaced with those of A ,

$$\lambda_1(A) > \lambda_1(M) \geq \lambda_2(A) \geq \lambda_2(M) \geq \dots \geq \lambda_N(A) \geq \lambda_N(M).$$

Hence, increasing $\lambda_2(A)$ implies increasing $\lambda_1(M)$. For regular graphs, all eigenvalues of the modularity matrix M are the same as those of the adjacency matrix A , except that $\lambda_1(A)$ is replaced with a zero eigenvalue.

C. Maximizing the modularity

Maximizing the modularity for $c > 2$ is a non-polynomial (NP)-complete problem [21,23]. However, Stoer and Wagner [24] presented a highly efficient algorithm to compute the minimum-cut problem when $c=2$, such that the modularity clustering for $c=2$ is not NP hard. In this section, we consider the spectral form (13) to deduce further insight along the lines of Newman in [25]. We define the non-negative weights

$$v_j = \sum_{k=1}^c (w_j^T s_k)^2,$$

and the modularity in Eq. (13) becomes

$$m = \frac{1}{2L} \sum_{j=1}^N v_j \lambda_j(M). \quad (19)$$

First, assume that $v_1 = \sum_{k=1}^c (w_1^T s_k)^2 = N$, then Eq. (16) implies that $w_j^T s_k = 0$ for all $1 \leq k \leq c$ and all $j > 1$. This means that the vectors s_1, s_2, \dots, s_c (or any linear combination of them, apart from $\sum_{k=1}^c s_k = u$) are orthogonal to the eigenvectors w_2, w_3, \dots, w_N . Since eigenvectors span the

N -dimensional space, it means that all s_k 's must be parallel or proportional to w_1 . However, the vectors s_1, s_2, \dots, s_c are orthogonal; hence, this is not possible.

If there are c clusters, it seems that we must require in Eq. (16) that

$$\sum_{j=1}^r v_j = N,$$

for $r \geq c-1$, and that necessarily at least $c-1$ eigenvalues in Eq. (13) play a role, because then $w_j^T s_k = 0$ for all $1 \leq k \leq c$ and all $j \geq c$. This means that the vectors s_1, s_2, \dots, s_c (or any linear combination of them, apart from $\sum_{k=1}^c s_k = u$ that is proportional to $w_q = \frac{u}{\sqrt{N}}$) are orthogonal to the eigenvectors w_c, w_{c+1}, \dots, w_N or, equivalently, that each community vector s_k (with 0 or 1 components) is a linear combination of the first $c-1$ eigenvectors of M , $s_k = \sum_{j=1}^{c-1} \beta_{kj} w_j$, where $\beta_{kj} = w_j^T s_k = \sum_{q \in C_k} (w_j)_q$. Together with the equation $\sum_{k=1}^c s_k = u$, this set of equations is sufficient to determine all c community vectors. How to choose these coefficients β_{kj} to optimize Eq. (13) remains a difficult problem. For example, we may adopt the strategy to choose the weights v_j , so that $v_j \geq v_{j+1}$ for each $1 \leq j < c$. However, by incorporating an additional eigenvector w_c , it can be possible to increase the weight v_1 corresponding to $\lambda_1(M)$ more [despite a lower $\lambda_c(M)$ is included in sum (13) and an extra v_c in $\sum_{j=1}^c v_j = N$, such that the average weight $E[v] = \frac{N}{c}$ decreases]. Numerical computations (see Fig. 4 below) show that all vector components in Eq. (13) seem to play a role, but that the first eigenvector w_1 is by far the most important.

For a regular graph with degree r , it is known [22] that $\lambda_1(M) = \lambda_2(A) < \lambda_1(A) = r = E[D]$ such that bound (18) equals

$$m_{\text{regular graph}} \leq \frac{\lambda_2(A)}{\lambda_1(A)} \left(1 - \frac{1}{c}\right) < 1 - \frac{1}{c}.$$

Since in general $\lambda_1(A) > \lambda_1(M) \geq \lambda_2(A)$, we observe that the lowest upper bound in Eq. (18) is reached for regular graphs. Another consequence of the interlacing is that in graphs with large spectral gap $\lambda_1(A) - \lambda_2(A)$, the largest eigenvalue $\lambda_1(M)$ can be much smaller than $\lambda_1(A)$. For example, the complete graph that possesses the largest possible spectral gap equal to N has $\lambda_1(M) = 0$, the lowest possible largest eigenvalue of any modularity matrix M . Intuitively, graphs with large spectral gap are difficult to tear apart, which means that they form already a quite tight community or cluster and that further dividing such a graph is hardly possible, resulting in a low modularity m .

III. ASSORTATIVITY

Networks where high-degree nodes preferentially connect to other high-degree nodes are called *assortative*, whereas networks where high-degree nodes connect to low-degree nodes are called *disassortative*. Assortativity is measured by the linear degree correlation coefficient ρ_D , but we use here assortativity and ρ_D interchangeably.

In [26], it has been shown that increasing the assortativity also increases the lower bound for $\lambda_1(A)$, but not necessarily

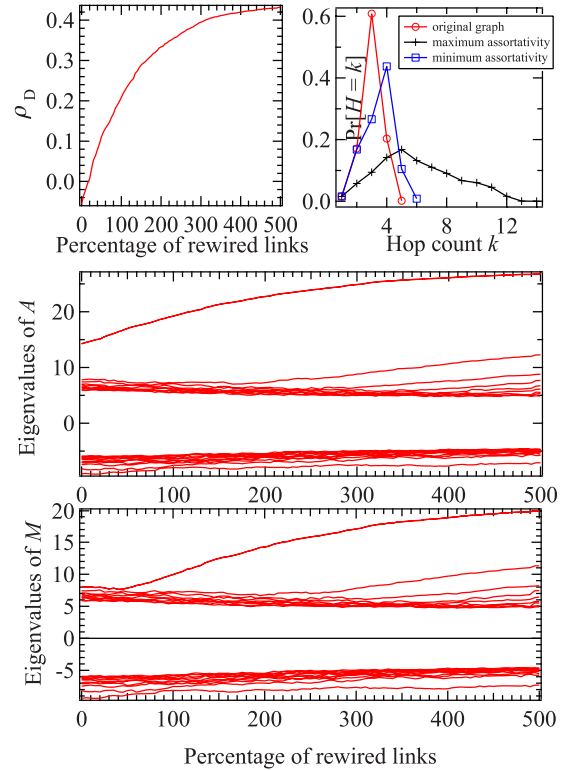


FIG. 1. (Color online) The 20 largest (in absolute value) eigenvalues of the adjacency matrix A and of the modularity matrix M as function of the percentage degree-preserving rewired links in an instance of the Barabási-Albert scale-free graph with $N=500$ and $L=1960$ links. The relation between ρ_D and the percentage of rewired links as well as three hop count distributions is also plotted.

$\lambda_2(A)$. However, by degree-preserving rewiring the matrix $\frac{1}{2L} d \cdot d^T$ is not changed, only A is. This implies that increasing ρ_D via degree-preserving rewiring does not change the sum of the eigenvalues of M , but it may increase the upper bound of $\lambda_1(M)$ and, hence, via Eq. (18) also the modularity. In any case, it will not decrease the upper bound for $\lambda_1(M)$, as follows from the interlacing property above.

A. Assortativity, maximum modularity, and the spectrum of A and M

Figure 1 shows the 20 largest (in absolute value) eigenvalues of both the adjacency matrix A and the modularity matrix M of a realization of the Barabási-Albert scale-free graph with $N=500$ nodes, $L=1960$ links, $E[D] \approx 7.85$, and $\rho_D \approx -0.05$. Each elementary degree-preserving rewiring step,³ specified by the lemma in [26] that changes the assortativity, results in a connected different graph (with the same degree vector). From one rewiring step to another, the large majority of eigenvalues of A (and, similarly, of M) are inter-

³A degree-preserved rewired Barabási-Albert scale-free graph (and similarly an Erdős-Rényi random graph), where ρ_D is significantly changed, is not a Barabási-Albert scale-free graph anymore, which is characterized by $\rho_D \rightarrow 0$ asymptotically (and $\rho_D = 0$ for Erdős-Rényi random graphs as shown in [26]).

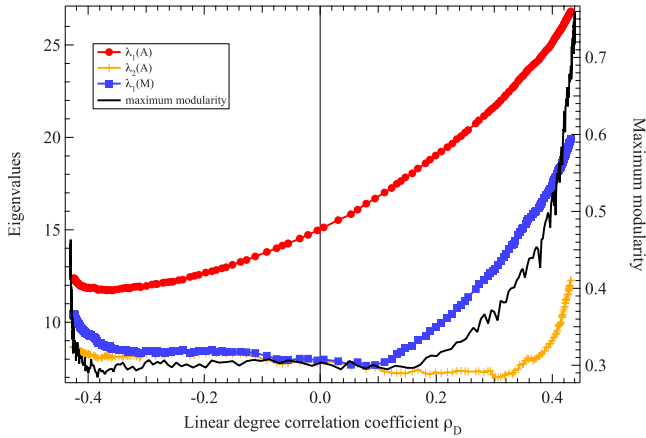


FIG. 2. (Color online) The largest and second largest eigenvalues of the adjacency matrix A and the largest eigenvalues of the modularity matrix M versus the linear degree correlation coefficient ρ_D for the Barabási-Albert graph with $N=500$ nodes and $L=1960$ links. The right-hand side axis shows the corresponding maximum modularity.

laced as shown in [26], while always, the eigenvalues of M are interlaced by those of A . The white band of eigenvalues around zero in Fig. 1 thus contains 480 smaller eigenvalues (which are not shown because the picture would color completely). Also the relationship between assortativity (via ρ_D) and the percentage of degree-preserving rewired links is shown, together with the hop count distribution of the original graph (no rewiring), and that at ρ_{Dmax} and ρ_{Dmin} .

Figures 2 and 3 show, for a realization of the rewired Barabási-Albert scale-free graph and of the rewired Erdős-Rényi random graph with equal N and almost equal number of links L , respectively, how the first few eigenvalues of the adjacency and modularity matrix vary with the linear degree

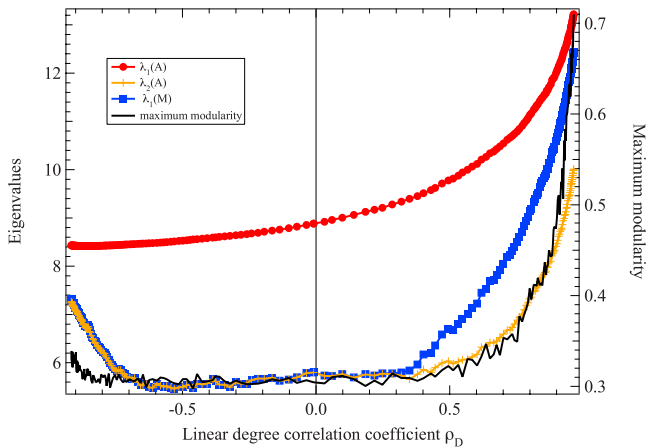


FIG. 3. (Color online) The largest and second largest eigenvalues of the adjacency matrix A and the largest eigenvalues of the modularity matrix M versus the linear degree correlation coefficient ρ_D for the Erdős-Rényi random graph $G_p(N)$ with $N=500$ nodes $L=1955$ links and $\rho_D \approx -0.01$. Thus, the link density $p=L/\binom{N}{2}$ equals $p \approx 1.25p_c$, where $p_c \sim \frac{\log N}{N}$ is the critical disconnectivity threshold. The right-hand side axis shows the corresponding maximum modularity.

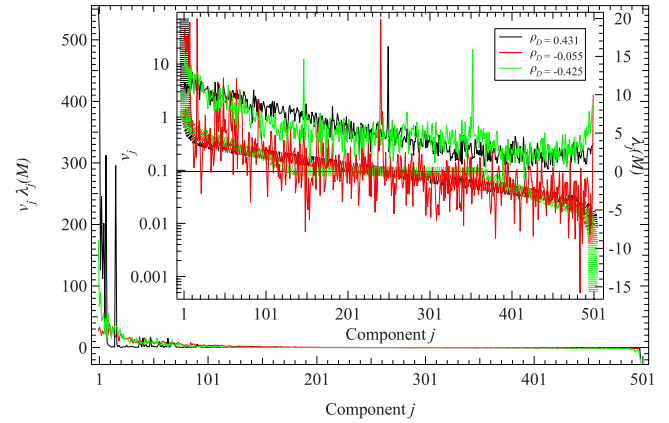


FIG. 4. (Color online) The product $v_j \lambda_j(M)$ for each component j for three instances: the original BA graph (red), the maximum assortativity rewired version (black), and the maximum disassortative rewired graph (green). The sum over all j equals $2Lm$ according to Eq. (13). The inset shows the weights v_j (at the left on logarithmic scale) and the eigenvalues $\lambda_j(M)$ (at the right in bold dashed lines).

correlation coefficient ρ_D . We observe for the two realizations of different classes of graphs that, in the disassortativity region ($\rho_D < 0$), $\lambda_1(M)$ follows $\lambda_2(A)$ reasonably well, while in the assortativity region $\lambda_1(M)$ starts increasing toward $\lambda_1(A)$. However, $\lambda_1(M)$ never reaches $\lambda_1(A)$, because $\lambda_1(M)$ is always strictly smaller than $\lambda_1(A)$, as proved in [22]. As a consequence and assuming that there is room to increase or decrease ρ_D , the larger the spectral gap, $\lambda_1(A) - \lambda_2(A)$, the larger the potential increase in modularity that can be achieved via degree-preserving rewiring.

Figures 2 and 3 also illustrate that the maximum modularity is roughly proportional to $\lambda_1(M)$ as long as $\lambda_1(M)$ is close to $\lambda_2(A)$. The maximum modularity has been computed by the approximate algorithm described in [27]. For increasing assortativity, the maximum modularity seems to increase faster than $\lambda_1(M)$. Apart from the extent in assortativity range, the rewired Barabási-Albert scale-free graph (Fig. 2) and the rewired Erdős-Rényi random graph (Fig. 3), both with same number of nodes and almost same number of links, behave surprisingly similarly, in spite of their different degree vectors.

For three instances of the rewired Barabási-Albert scale-free graph, Fig. 4 draws each term $v_j \lambda_j(M)$ in the spectral form (13) of the modularity, as well as each weight v_j and eigenvalue $\lambda_j(M)$ for $1 \leq j \leq 500$. The inset in Fig. 4 shows that the weights v_j vary irregularly, as a noisy signal around the mean 1, and that a very high peak (on a logarithmic scale) is observed corresponding to $\lambda_q(M)=0$, which has inspired us to the general bound (18). Apart from that peak corresponding to the eigenvector $w_q = \frac{u}{\sqrt{N}}$, the weights roughly decrease with the component or eigenvector j . Apart from a few values, the resulting product $v_j \lambda_j(M)$ is decreasing in j . Although only shown for the Barabási-Albert scale-free graph, these observations are generally observed: the first term $v_1 \lambda_1(M)$ contributes dominantly to the modularity (19) and illustrates that bound (18) can be sharp. The other terms

$j > 1$ are initially positive, but then negative (because the eigenvalues become negative), and the whole sum is needed to compute the modularity. Remarkably, a huge cancellation in the sum occurs because we found that sum (13) is close to its first term.

“Shifting-the-weights” principle

Since the eigenvalues of M are ordered as usual, $\lambda_1(M) \geq \lambda_2(M) \geq \dots \geq \lambda_N(M)$, the maximum modularity is achieved by shifting in Eq. (13) as much weight as possible to the larger eigenvalues, which we call the shifting-the-weights principle. Figure 4 supports this principle: fewer eigenvalues in Eq. (13) imply that the individual weights v_j are higher on average due to condition (16). Furthermore, Fig. 4 illustrates that, especially in the high assortativity regime, the first c eigenvalues are clearly dominant, as argued in Sec. II C.

When the largest eigenvalues are close to each other, incorporating additional eigenvalues may increase the weights on the largest eigenvalues, which leads to a larger modularity. As the assortativity increases, the largest eigenvalues of matrix M seem to be dispelled from each other (see Fig. 1). In other words, spacing $\lambda_1(M) - \lambda_2(M), \lambda_2(M) - \lambda_3(M), \dots$ between the largest eigenvalues of the matrix M seems to grow as the assortativity increases. For large ρ_D , the maximal modularity includes a minimum amount⁴ of eigenvalues in Eq. (13). The chance to increase the modularity m by incorporating more eigenvalues is small because (a) the average weight is reduced due to condition (16) and (b) the extra eigenvalues included are far smaller when the spacings of the leading eigenvalues are large. As a result, the modularity increases with increasing assortativity ρ_D faster than $\lambda_1(M)$ as shown in Figs. 2 and 3, because v_1 also increases with ρ_D . Since the smallest number of eigenvalues that play a role in the maximum modularity is $c - 1$, the “shifting-the-weights” principle also implies that the number of clusters c decreases with increasing assortativity, which is also observed in Fig. 5.

B. Number and size of clusters when maximizing the modularity

Figure 5 illustrates that the number of clusters as a function of ρ_D in the rewired Erdős-Rényi random graph is roughly a stretch-out version of that in the rewired Barabási-Albert scale-free graph. We also observe that the number of clusters “inversely” correlates with the maximum modularity in Figs. 2 and 3: a high maximum modularity seems to correspond to a low number of clusters (and conversely). Finally, Fig. 5 also shows the cluster estimate $c^* = 2\sqrt{2L}/D_{\Delta C}$ in Eq. (7).

⁴This observation is consistent with what we can deduce from the modularity bound (A1) expressed in terms of the eigenvalue spacings: when the spacing is generally large, including an extra eigenvalue [i.e., $c \rightarrow c + 1$ in Eq. (A1) in Appendix A], will cause a decrease in the last term $N\lambda_{c+1}(M)$ of Eq. (A1), which is hardly exceeded by the increase in the first sum, due to the large weight N in that last term and condition (16).

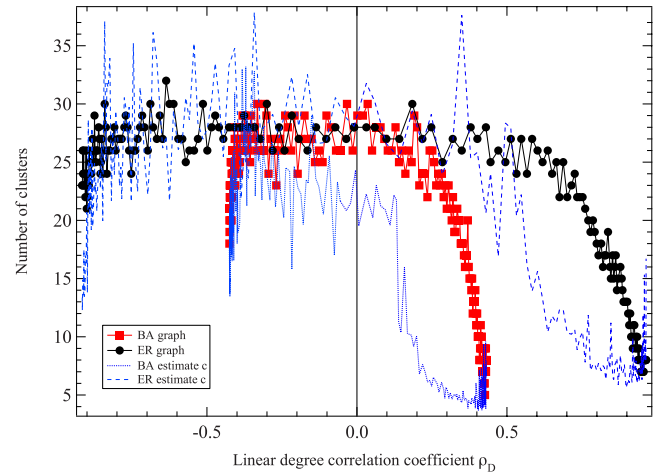


FIG. 5. (Color online) The number of clusters for both the Barabási-Albert scale-free graph (corresponding to Fig. 2) and the Erdős-Rényi random graph (corresponding to Fig. 3) as a function of the linear degree correlation coefficient ρ_D . For both graphs, also the estimate in Eq. (7) is drawn in dotted line.

Figures 6 and 7 show details per cluster of the rewired Barabási-Albert scale-free graph and Erdős-Rényi random graph for the extreme assortativity and disassortativity and two intermediate values of ρ_D . The cluster sizes are ranked from largest to smallest. Apart from the extreme assortative graphs, the cluster sizes are about the size order of magnitude, as well as the average degree of nodes in each cluster, which is drawn in the inset. The larger the assortativity, the fewer the number of clusters (Fig. 5) and the larger the differences in cluster size and average degree per cluster.

We mention that community sizes in real-world networks have been reported to follow a power law [5]. Our graph instances are far too small ($N = 500$) to see a power-law behavior. However, only in highly assortative graphs we expect a power law assuming that our results can be scaled to larger sizes of N . In disassortative graphs, the sizes of the clusters

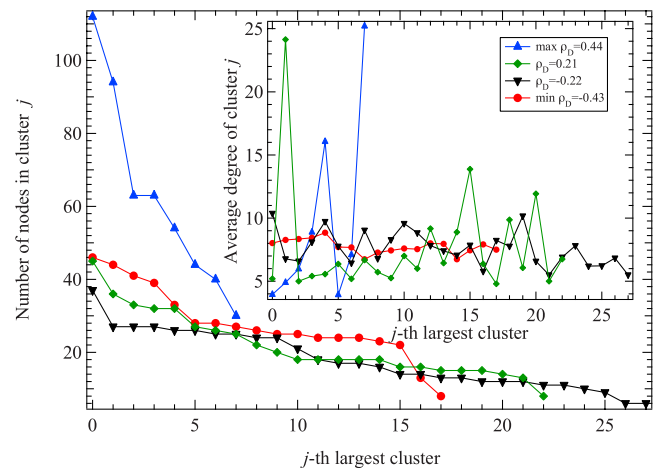


FIG. 6. (Color online) The cluster size of the j th largest cluster for four values of ρ_D of the Barabási-Albert scale-free graph (corresponding to Fig. 2). The inset shows the average degree for each cluster j .

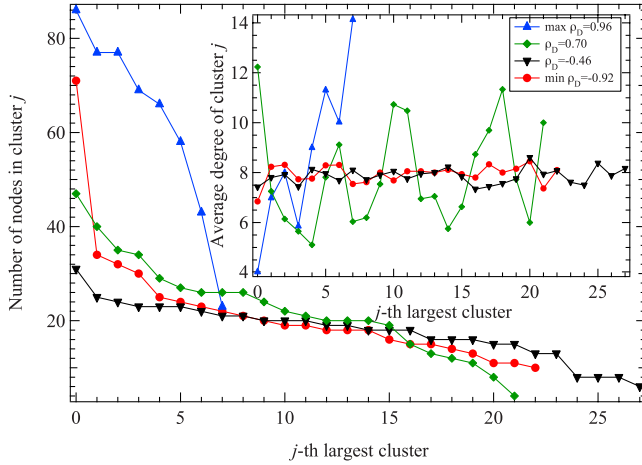


FIG. 7. (Color online) The cluster size of the j th largest cluster for four values of ρ_D of the Erdős-Rényi random graph (corresponding to Fig. 3). The inset shows the average degree for each cluster j .

of the maximized modularity are roughly the same, as well as the number of links (or average degree). For these graphs where $\text{Var}[n_C] \approx 0$, the spectral upper bound (18) seems promising.

C. Hop count and the assortativity

The hop count H is the number of links in an arbitrary shortest path in the graph. The average hop count $E[H]$ is a measure for the distance between nodes in the graph and reflects the efficiency of transport in the graph: a short hop count means that transport only requires a few links and that the end-to-end delay is likely small (provided it is a sum over the number of hops in a path).

Figure 8 shows how the average hop count $E[H]$ plus or minus one standard deviation $\sigma_H = \sqrt{\text{Var}[H]}$ changes with the

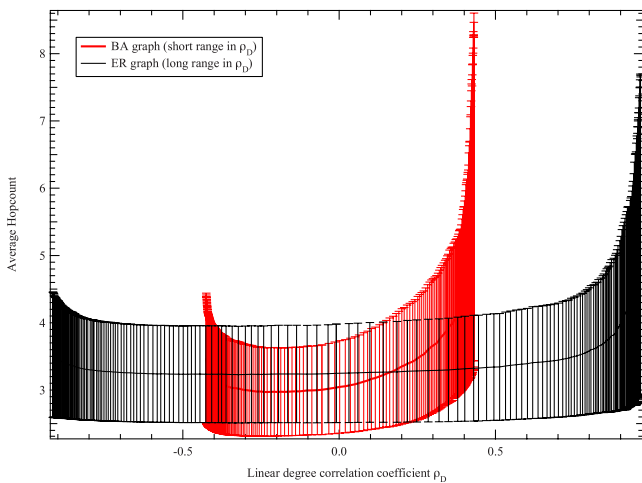


FIG. 8. (Color online) The average hop count $E[H]$ plus or minus one standard deviation $\sigma_H = \sqrt{\text{Var}[H]}$ for both the Barabási-Albert scale-free graph (corresponding to Fig. 2) and the Erdős-Rényi random graph (corresponding to Fig. 3) as a function of the linear degree correlation coefficient ρ_D .

linear degree correlation coefficient ρ_D . In both realizations, we observe from Fig. 8 that the standard deviation σ_H is smaller than the average, except for the extreme assortativity in the Barabási-Albert scale-free graph that has a broad bell-shaped hop count distribution (see Fig. 1). In addition, we observe that $E[H] + \sigma_D$ correlates well with the maximum modularity. An increase in average hop count $E[H]$ with increasing assortativity agrees with the increase in the largest eigenvalue $\lambda_1(A)$ of the adjacency matrix as function of ρ_D . For example, virus or information spread (see, e.g., [28]) is mainly possible when the effective spreading strength exceeds a threshold $\tau_c = 1/\lambda_1(A)$. Thus, when $\lambda_1(A)$ is large (corresponding to a large ρ_D), virus spread is easy, although the hop count is large, while the opposite holds for the disassortative region. The increase in the largest eigenvalue $\lambda_1(A)$ suggests that virus or information spread over a part of the network becomes easier. Actually, it is easy to spread over the high-degree nodes which are well interconnected in an assortative network. However, the spread over the entire network, especially over the low-degree nodes can be difficult, as implied by the increased average hop count.

For sparse large graphs, the average hop count is approximately (see [29], pp. 343–345)

$$E[H] \approx \frac{\log N}{\log \mu},$$

where μ is the average degree minus 1. Assuming that the number of nodes of each cluster is more or less the same (see Figs. 6 and 7), the average hop count in a modular or hierarchical graph approximately consists of the average number of intercommunity hops multiplied by the average number of intracommunity hops,

$$E[H] \approx \frac{\log c}{\log \mu_{\text{inter}}} \frac{\log N/c}{\log \mu_{\text{intra}}},$$

from which we deduce that the minimum average hop count occurs, ignoring the integer nature of c , when $c = \sqrt{N}$. This asymptotic and approximate estimate agrees reasonably well with Figs. 5 and 8, where $\sqrt{N} \approx 22$. This square-root law $c = O(\sqrt{N})$ is remarkable, because earlier Fortunato [10] has constructed an example that maximizes the modularity from which he found that $c = \sqrt{L}$. We can indeed show, more generally, that a maximization of the modularity with respect to the number of clusters c results in $c = \sqrt{L} + r$, where the smaller terms are $r = O(1)$. Hence, when the modularity is maximized for sufficiently large graphs, it seems likely that the size of the clusters scales as $O(\sqrt{N})$, whereas the number of links in a cluster scales as $O(\sqrt{L})$, and that this holds for sparse graphs where $L = O(N)$, as follows from $c = O(\sqrt{N}) = \sqrt{L}$.

For both graphs, $E[H]$ increases as the assortative modules are forming in the graphs. Within the modules, the hop counts are lowered, but between the modules, the hop counts are increased, resulting in the observed increase in variance. For the rewired Barabási-Albert scale-free graph, the grouping of the highest-degree nodes causes several hop counts to increase as the efficient transport (originally, $E[H] \approx 3$) that is associated with scale-free graphs is lost. Thus, although

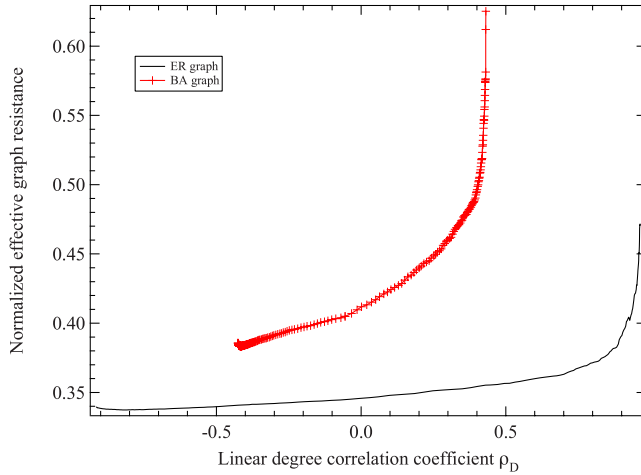


FIG. 9. (Color online) The normalized effective graph resistance $R_G/\binom{N}{2}$ for both the Barabási-Albert scale-free graph (corresponding to Fig. 2) and the Erdős-Rényi random graph (corresponding to Fig. 3) versus the linear degree correlation coefficient ρ_D .

the grouping of nodes with similar degrees does produce several short paths within the groups, the loss of the hubs as transport facilitators reduces the number of short paths. In spite of the large difference in the degree distributions (albeit the average degree $E[D]=\frac{2L}{N}$ is the same), Fig. 8 shows that the average hop count $E[H]$ plus or minus one standard deviation $\sigma_H=\sqrt{\text{Var}[H]}$ is surprisingly similar in these both graphs.

D. Effective graph resistance

The effective graph resistance R_G is defined (see, e.g., [22]) by

$$R_G = N \sum_{k=1}^{N-1} \frac{1}{\mu_k},$$

where μ_k is the k th largest eigenvalue of the Laplacian of the graph. The effective graph resistance R_G measures the ease of communication in a graph. It can be shown [22] that

$$E[H] \geq \frac{R_G}{\binom{N}{2}}.$$

Figure 9 shows the normalized effective graph resistance $R_G/\binom{N}{2}$ as a function of the assortativity. While the average hop count of the rewired Barabási-Albert scale-free graph can be smaller than that of the rewired Erdős-Rényi random graph (Fig. 8), its corresponding effective graph resistance is always larger.

Just as the average hop count $E[H]$, the normalized effective graph resistance $R_G/\binom{N}{2}$ of a graph seems to correlate highly with the modularity as a function of ρ_D . In addition, the sharp increase toward the maximum value of ρ_D indicates that disconnectivity (characterized by $R_G \rightarrow \infty$) is likely to occur when ρ_D would be further increased.

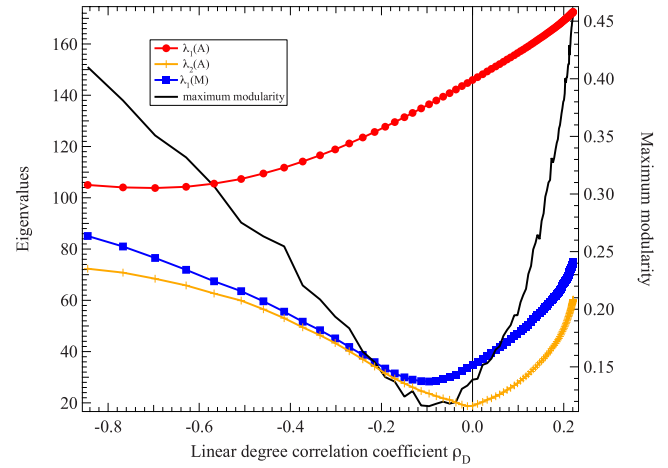


FIG. 10. (Color online) The largest and second largest eigenvalues of the adjacency matrix A and the largest eigenvalues of the modularity matrix M versus the linear degree correlation coefficient ρ_D for the air transportation network with $N=2179$ nodes, $L=31\,326$ links, and $E[D]\approx 28.8$. The right-hand side axis shows the corresponding maximum modularity.

IV. REAL-WORLD GRAPHS

Figures 10–13 illustrate that similar phenomena, deduced above from instances of the Erdős-Rényi random graph and the Barabási-Albert scale-free graph, also occur in real-world networks. The details of these networks are found for the air transportation network of the USA in [26], for the peer-to-peer network Gnutella in [30], for the coappearance network of characters in the novel *Les Misérables* of Victor Hugo in [31], and for the protein residue network of the immunoglobulin 1A4J in [32].

We observe the general trend that (a) the largest eigenvalue $\lambda_1(M)$ of the modularity matrix is close to $\lambda_2(A)$, except for extreme values of ρ_D , and (b) the maximum modu-

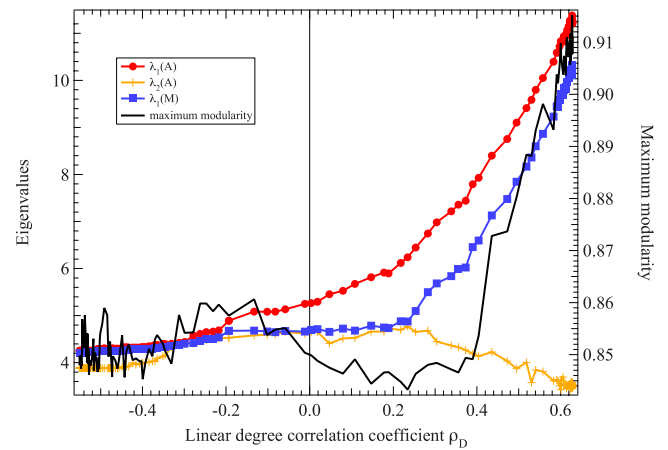


FIG. 11. (Color online) The largest and second largest eigenvalues of the adjacency matrix A and the largest eigenvalues of the modularity matrix M versus the linear degree correlation coefficient ρ_D for the Gnutella network with $N=77$ nodes, $L=254$ links, and $E[D]\approx 6.5$. The right-hand side axis shows the corresponding maximum modularity.

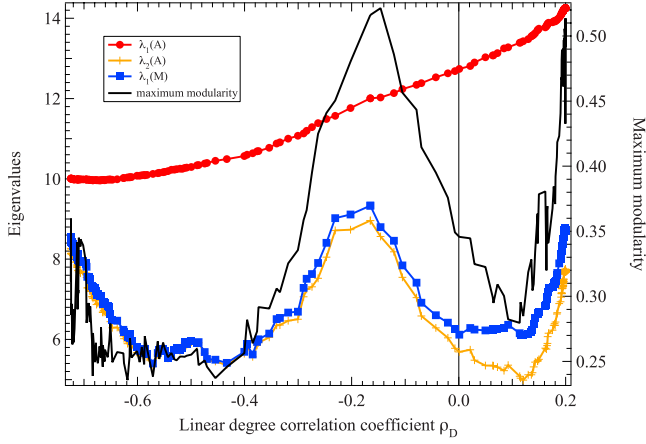


FIG. 12. (Color online) The largest and second largest eigenvalues of the adjacency matrix A and the largest eigenvalues of the modularity matrix M versus the linear degree correlation coefficient ρ_D for the network Les Miserables with $N=737$ nodes, $L=803$ links, and $E[D] \approx 2.2$. The right-hand side axis shows the corresponding maximum modularity.

larity increases with sufficiently large ρ_D and faster than $\lambda_1(M)$. That the trend is not universal is demonstrated in Fig. 13, where part (b) of the trend is not followed. This observed trend restricts the hypothesis in [7] that modularity seems to increase in evolutionary processes; it mainly does if these processes are assortative in nature. As explained in [26], degree-preserving rewiring may be regarded as a (simplified) evolutionary process and most complex networks are not assortative (which is in agreement with the data in [15]), because highly assortative networks are likely disconnected and lack synergy or diversity in their connection pattern.

V. CONCLUSION

Several expressions (3), (9), and (10) and bounds (4), (5), and (18) for Newman's modularity are derived. These results

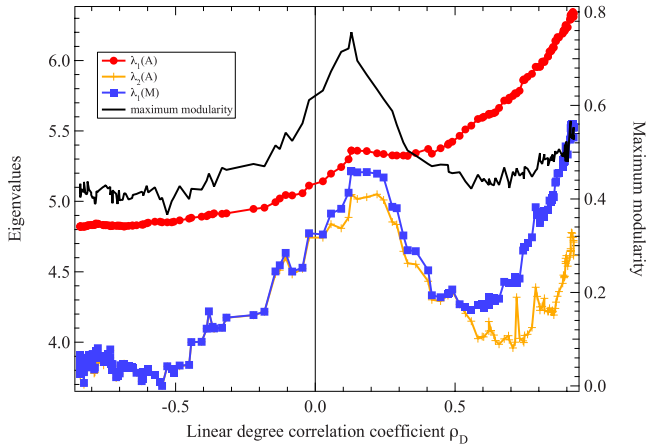


FIG. 13. (Color online) The largest and second largest eigenvalues of the adjacency matrix A and the largest eigenvalues of the modularity matrix M versus the linear degree correlation coefficient ρ_D for the protein network of immunoglobulin 1A4J with $N=95$ nodes, $L=213$ links, and $E[D] \approx 4.5$. The right-hand side axis shows the corresponding maximum modularity.

reveal conditions for or properties of the maximum modularity of a network and provide proofs of earlier claims. The influence of the spectrum of the modularity matrix on the maximum modularity is discussed from which we deduce a spectral bound (18).

The second part of the paper investigates how the maximum modularity, the number of clusters, and the hop count of the shortest paths vary when the assortativity of the graph is changed via degree-preserving rewiring. Simulations suggest that, apart from the heavy assortative regime, the maximum modularity behaves as $\lambda_1(M)$, which—in turn—closely follows the second largest eigenvalue $\lambda_2(A)$ of the adjacency matrix A . Extensive simulations on several real-world complex networks show that the maximum modularity increases, the number of clusters decreases, and the average hop count and the effective graph resistance increase with increasing assortativity. The major driver to study the influence of degree-preserved rewiring on modularity is to shed light on evolutionary processes in nature. If degree-preserved rewiring can be regarded as a simplified model for an evolutionary process, then we show that networks, close to their assortative bounds, generally possess a clear modular structure.

ACKNOWLEDGMENT

The work of P.S. was supported by the National Agricultural Biosecurity Center (NABC) at Kansas State University.

APPENDIX A: ALTERNATIVE FORM OF THE MODULARITY IN TERMS OF THE EIGENVALUE SPACING

The partial (or Abel) summation,

$$\sum_{k=1}^n a_k b_k = \sum_{k=1}^{n-1} \left(\sum_{l=1}^k a_l \right) (b_k - b_{k+1}) + b_n \left(\sum_{l=1}^n a_l \right),$$

of $m^* = (2L)m = \sum_{j=1}^N v_j \lambda_j(M)$ equals

$$\begin{aligned} m^* &= \sum_{j=1}^{N-1} \left(\sum_{k=1}^j v_k \right) [\lambda_j(M) - \lambda_{j+1}(M)] + \lambda_N(M) \sum_{k=1}^N v_k \\ &= \sum_{j=1}^{N-1} \left(\sum_{k=1}^j v_k \right) [\lambda_j(M) - \lambda_{j+1}(M)] + N\lambda_N(M), \end{aligned}$$

and only the last term is negative, while all others are positive. Further, using condition (16) such that $\sum_{k=1}^j v_k + \sum_{k=j+1}^N v_k = N$, we can write

$$\begin{aligned} m^* &= \sum_{j=1}^c \left(\sum_{k=1}^j v_k \right) [\lambda_j(M) - \lambda_{j+1}(M)] + \sum_{j=c+1}^{N-1} \left(N - \sum_{k=j+1}^N v_k \right) \\ &\quad \times [\lambda_j(M) - \lambda_{j+1}(M)] + N\lambda_N(M) \\ &= \sum_{j=1}^c \left(\sum_{k=1}^j v_k \right) [\lambda_j(M) - \lambda_{j+1}(M)] + N\lambda_{c+1}(M) \\ &\quad - \sum_{j=c+1}^{N-1} \left(\sum_{k=j+1}^N v_k \right) [\lambda_j(M) - \lambda_{j+1}(M)]. \end{aligned}$$

Again, by shifting all the weight to the first c clusters such that $\lambda_{c+1}(M) \geq 0$ and $v_k = 0$ for $k \geq c+1$, an upper bound of the modularity can be written in terms of the spacings $\lambda_j(M) - \lambda_{j+1}(M)$ as

$$m^* \leq \sum_{j=1}^c \left(\sum_{k=1}^j v_k \right) [\lambda_j(M) - \lambda_{j+1}(M)] + N\lambda_{c+1}(M). \quad (\text{A1})$$

This expression shows that the larger indices of j are more heavily weighted. Recall that the spacing $\Delta\lambda_j = \lambda_j(M) - \lambda_{j+1}(M) \geq 0$ is not necessarily decreasing with j .

The number of positive eigenvalues of M is, due to interlacing, less (by 1 or 0) than the number of positive eigenvalues of A . The latter is larger than the independence number (see [22]), which is equal to the largest coclique. Thus, when maximizing the modularity [such that $\lambda_{c+1}(M) \geq 0$], the number of clusters c should be smaller than the independence number of the graph.

APPENDIX B: ALTERNATIVE MATRIX FORM FOR THE MODULARITY

Identities (8) and (9) also offer a matrix representation of the modularity as

$$m = \frac{1}{(2L)^2} u^T [D_C D_C^T - \text{diag}(D_C^2)] u - \frac{L_{\text{inter}}}{L}, \quad (\text{B1})$$

where D_C is the $c \times 1$ vector of community degree sums, which equals

$$D_C = S^T d.$$

Using that relation gives

$$\begin{aligned} m &= \frac{1}{(2L)^2} u^T D_C D_C^T u - \frac{1}{(2L)^2} D_C^T D_C - \frac{L_{\text{inter}}}{L} \\ &= \frac{1}{(2L)^2} u^T S^T d d^T S u - \frac{1}{(2L)^2} d^T S S^T d - \frac{L_{\text{inter}}}{L}. \end{aligned}$$

With $S u_{c \times 1} = u_{N \times 1}$ and $u^T d = 2L$, we arrive at

$$m = 1 - \frac{L_{\text{inter}}}{L} - \frac{1}{(2L)^2} d^T S S^T d,$$

which equals Eq. (2).

-
- [1] X. Chu, J. Guan, Z. Zhang, and S. Zhou, *J. Stat. Mech.: Theory Exp.* (2009), P07043.
- [2] J. P. Gleeson, *Phys. Rev. E* **77**, 046117 (2008).
- [3] J. Omic, J. Martin Hernandez, and P. Van Mieghem, 22nd International Teletraffic Congress (ITC 22), Amsterdam, (2010).
- [4] M. E. J. Newman and M. Girvan, *Phys. Rev. E* **69**, 026113 (2004).
- [5] S. Fortunato, *Phys. Rep.* **486**, 75 (2010).
- [6] I. A. Kovács, R. Palotai, M. S. Szalay, and P. Csermely, *PLoS One* **5**, e12528 (2010).
- [7] R. Guimerà, M. Sales-Pardo, and L. A. Nunes Amaral, *Phys. Rev. E* **70**, 025101(R) (2004).
- [8] L. Danon, A. Díaz-Guilera, J. Duch, and A. Arenas, *J. Stat. Mech.: Theory Exp.* (2005), P09008.
- [9] J. Reichardt and S. Bornholdt, *Phys. Rev. E* **74**, 016110 (2006).
- [10] S. Fortunato and M. Barthélemy, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 36 (2007).
- [11] J. Kumpula, J. Saramäki, K. Kaski, and J. Kertész, *Eur. Phys. J. B* **56**, 41 (2007).
- [12] A. Arenas, A. Fernández, and S. Gómez, *New J. Phys.* **10**, 053039 (2008).
- [13] B. H. Good, Y. de Montjoye, and A. Clauset, *Phys. Rev. E* **81**, 046106 (2010).
- [14] M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).
- [15] M. E. J. Newman, *Phys. Rev. E* **67**, 026126 (2003).
- [16] M. E. J. Newman and M. Girvan, *Mixing Patterns and Community Structure in Networks*, Statistical Mechanics of Complex Networks (Springer, Berlin, 2003), pp. 66–87.
- [17] R. Xulvi-Brunet and I. Sokolov, *Phys. Rev. E* **70**, 066102 (2004).
- [18] M. Brede and S. Sinha, e-print [arXiv:cond-mat/0507709](https://arxiv.org/abs/cond-mat/0507709).
- [19] P. Holme and J. Zhao, *Phys. Rev. E* **75**, 046111 (2007).
- [20] Z. Jing, T. Lin, Y. Hong, L. Jian-Hua, C. Zhi-Wei, and L. Yi-Xue, *Chin. Phys.* **16**, 3571 (2007).
- [21] M. E. J. Newman, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8577 (2006).
- [22] P. Van Mieghem, *Graph Spectra for Complex Networks* (Cambridge University Press, Cambridge, UK) (2010).
- [23] U. Brandes, D. Delling, M. Gaertler, R. Görke, M. Hofer, Z. Nikoloski, and D. Wagner, *IEEE Trans. Knowl. Data Eng.* **20**, 172 (2008).
- [24] M. Stoer and F. Wagner, *J. ACM* **44**, 585 (1997).
- [25] M. E. J. Newman, *Phys. Rev. E* **74**, 036104 (2006).
- [26] P. Van Mieghem, H. Wang, X. Ge, S. Tang, and F. A. Kuipers, *Eur. Phys. J. B* **76**, 643 (2010).
- [27] A. Clauset, M. E. J. Newman, and C. Moore, *Phys. Rev. E* **70**, 066111 (2004).
- [28] P. Van Mieghem, J. Omic, and R. E. Kooij, *IEEE/ACM Trans. Netw.* **17**, 1 (2009).
- [29] P. Van Mieghem, *Performance Analysis of Communications Systems and Networks* (Cambridge University Press, Cambridge, UK, 2006).
- [30] Gnutella: Internet Peer-to-Peer Networks, <http://snap.stanford.edu/data/>
- [31] D. E. Knuth, *The Stanford GraphBase: A Platform for Combinatorial Computing* (Addison-Wesley, Reading, MA, 1993).
- [32] Protein Data Bank (PDB) Codes, <http://www.pdb.org/pdb/explore/explore.do?structureId=1A4J>
- [33] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, New York, 2010).