Linear Algebra and its Applications 473 (2015) 359–376

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Spectral properties of modularity matrices

Marianna Bolla [a,*], Brian Bullins [b], Sorathan Chaturapruek [c], Shiwen Chen [d], Katalin Friedl [e]

[a] *Institute of Mathematics, Budapest University of Technology and Economics and Center for Telecommunication and Informatics, Debrecen University, Hungary*
[b] *Dept. of Computer Science, Princeton University, Princeton, NJ, USA*
[c] *Dept. of Computer Science, Stanford University, Stanford, CA, USA*
[d] *Williams College, Williamstown, MA, USA*
[e] *Dept. of Computer Science and Information Theory, Budapest University of Technology and Economics, Hungary*

ARTICLE INFO

Article history:
Received 30 September 2013
Accepted 27 October 2014
Available online 7 November 2014
Submitted by L.-H. Lim

MSC:
05C50
62H17

Keywords:
χ^2 test for independence
Modularity matrix
Spectral clustering
Multipartite graphs
Isoperimetric number
Maximal correlation

ABSTRACT

There is an exact relation between the spectra of modularity matrices introduced in social network analysis and the χ^2 statistic. We investigate a weighted graph with the main interest being when the hypothesis of independent attachment of the vertices is rejected, and we look for clusters of vertices with higher inter-cluster relations than expected under the hypothesis of independence. In this context, we give a sufficient condition for a weighted, and a sufficient and necessary condition for an unweighted graph to have at least one positive eigenvalue in its modularity or normalized modularity spectrum, which guarantees a community structure with more than one cluster. This property has important implications for the isoperimetric inequality, the symmetric maximal correlation, and the Newman–Girvan modularity.

* Corresponding author.
E-mail addresses: marib@math.bme.hu (M. Bolla), bbullins@cs.princeton.edu (B. Bullins), tummykung@gmail.com (S. Chaturapruek), s.heidi.chen@gmail.com (S. Chen), friedl@cs.bme.hu (K. Friedl).

http://dx.doi.org/10.1016/j.laa.2014.10.039
0024-3795/

1. Introduction

When performing correspondence analysis on a *contingency table*, we use the singular value decomposition of the normalized table. The sum of the squares of these singular values (apart from the trivial 1) multiplied with the sample size (which is lost during the normalization) gives the value of the χ^2 statistic for testing *independence* of the two underlying categorical variables, see [4]. A symmetric table of zero diagonal corresponds to an undirected *weighted graph*, where the non-negative entries are pairwise similarities between the vertices. In [3] we introduced the *normalized modularity matrix* which is, in fact, the matrix of the normalized symmetric table deprived of the trivial factor. The spectral norm of this matrix bounds the discrepancy of the underlying weighted graph in the sense of the Expander Mixing Lemma, see [5,7]. By this lemma, a 'small' spectral norm indicates that the actual connections between the vertices are close to what is expected under independent vertex attachment, a fact that is in accord with a 'small' value of the χ^2 statistic. We are interested in the case of a relatively 'high' spectral norm and χ^2 value, when the number and the sign of the structural eigenvalues (separated from zero), together with some classification properties of the corresponding eigenvectors, will determine the so-called community (or anti-community) structure of the graph, see [3]. More precisely, based on the spectra we can count the clusters and, using the eigenvectors, find the clusters themselves, so that the intercluster relations within the clusters are higher (or lower) than expected under the null-model of independence.

The unnormalized version of the modularity matrix was introduced in [8] for the purposes of social network analysis. In [8,9] Newman and Girvan also defined their modularity as a nonparametric statistic, akin to the normalized or multiway cuts, measuring the community structure in a network. For a given positive integer k, the k-way *Newman–Girvan modularity* favors k-partitions of the vertices into k disjoint clusters (modules) within which the actual connections are higher than expected in a random graph. Even for given k, the optimal k-partition cannot be found in polynomial time in the number of vertices. One possibility is to first divide the vertices into two clusters, and if the two-way modularity of the optimal two-partition is positive, then further divide the clusters. However, it sometimes happens that the maximal two-way Newman–Girvan modularity is negative, and in this case there is no use of looking for further modules; the network is called *indivisible*, see [9]. In [10] a spectral method is proposed to divide the vertices into two parts based on the signs of the coordinates of the eigenvector corresponding to the largest eigenvalue of the modularity matrix. It can be done only if this eigenvalue is strictly positive, so the corresponding eigenvector has both positive and negative coordinates, being orthogonal to the all 1's vector (eigenvector belonging to the zero eigenvalue). On the contrary, if the largest eigenvalue is zero (zero is inevitably an eigenvalue), the network is indivisible in the above sense. We will show that having zero as the largest eigenvalue of the modularity matrix is a sufficient but not necessary condition for the network to be indivisible.

When the network is not indivisible, we can compute the modularity matrices of the subgraphs induced by the clusters of vertices giving the optimal two-partition and investigate their largest eigenvalues. If one of them is zero, that component is indivisible; otherwise, we proceed with finding the two-partition of it in the same way, and so forth. If we only want to test the largest eigenvalue of the subgraph for zero, then it suffices to find certain patterns in it, as we will see. Sometimes we are looking for a balanced partition of the vertices (with similar sizes or volumes), for which purpose the eigendecomposition of the normalized modularity matrix is used, see [3]. In [3] we also proved that the existence of $k-1$ positive eigenvalues of this matrix is an indication of a k-module structure and to approximate it we proposed the weighted k-means algorithm for the vertex representatives based on the corresponding eigenvectors. However, when the network is large, it suffices to use the upper end of the positive modularity spectrum separated from the bulk of the eigenvalues.

Since the allocation of the positive modularity eigenvalues plays an important role in the community detection problem, first we have to clarify when there are positive eigenvalues at all. In some examples of [3] we discussed that the modularity spectra of complete and complete multipartite graphs have no positive eigenvalues, exhibiting a one-module and anti-community structure, respectively. In the present paper, we will prove that both the modularity and the normalized modularity matrix of an unweighted graph is negative semidefinite if and only if the graph is complete or complete multipartite. We will also extend the above notions to weighted graphs and call a weighted graph *soft-core multipartite* if there is a partition of its vertices into clusters with edges of zero weight within, and positive weight between them. It is proved that whenever our graph is not soft-core multipartite, its modularity and normalized modularity matrix has a positive eigenvalue. The proof relies on the appearance of a special triangle in the graph, one that is quite common in real-life networks. The above characterization has many important implications for the symmetric *maximal correlation* and the *isoperimetric inequality*, see [2].

The paper is organized as follows. In Section 2 notation and some important facts about the modularity eigenvalues and the χ^2 statistic are introduced. In Section 3 modularity spectra of complete and complete multipartite graphs are derived. In Section 4 we state and prove the main results about the characterization of unweighted graphs with zero as the largest modularity eigenvalue and the extension to weighted graphs. In Section 5 we discuss some other applications of our main Theorems 10, 12, and 13, concerning the Newman–Girvan modularity, isoperimetric number, and maximal correlation.

2. Preliminaries

In [8] Newman and Girvan defined the *modularity matrix* of an unweighted graph on n vertices with the $n \times n$ symmetric adjacency matrix $\mathbf{A}$ as

$$\mathbf{M} = \mathbf{A} - \frac{1}{2e}\mathbf{d}\mathbf{d}^T, \tag{1}$$

where $\mathbf{d} = (d_1, \dots, d_n)^T$ is the so-called *degree-vector* comprised of the vertex-degrees d_i's and $2e = \sum_{i=1}^{n} d_i$ is twice the number of edges. We denote column vectors by lower-case bold letters; row vectors are written as transposes of column vectors. In [3] we formulated the *modularity matrix of a weighted graph* $G = (V, \mathbf{W})$ on the n-element vertex-set V with the $n \times n$ symmetric weight-matrix $\mathbf{W}$, the entries of which are pairwise similarities between the vertices and satisfy $w_{ij} = w_{ji} \geq 0$, $w_{ii} = 0$, as follows:

$$\mathbf{M} = \mathbf{W} - \mathbf{d}\mathbf{d}^T, \tag{2}$$

where the entries of $\mathbf{d}$ are the *generalized vertex-degrees* $d_i = \sum_{j=1}^{n} w_{ij}$ $(i = 1, \dots, n)$. Here $\mathbf{W}$ is normalized in such a way that $\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} = 1$, an assumption that does not hurt the generality, but simplifies further notation and makes it possible to consider $\mathbf{W}$ as a symmetric joint distribution of two identically distributed discrete random variables taking on n different values. However, with this normalization, we lose the sample size N in the case when we start with a symmetric contingency table of upper-diagonal counts corresponding to pairwise relations between the vertices. For example, the vertices may correspond to Facebook users and the counts indicate the number of pairwise communications between them; or the vertices may correspond to synopses of the brain and the counts indicate the number of signals transmitted between them (there are no self-communications). The sample size N is twice the sum of the counts, i.e., the total number of communications or transmissions. Then the null-hypothesis of $w_{ij} = d_i d_j$ (which means that the communications or transmissions happen independently, with probabilities proportional to the generalized degrees) is tested by the χ^2 statistic written in the following convenient form:

$$\chi^2 = N \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{(w_{ij} - d_i d_j)^2}{d_i d_j} = N \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{m_{ij}^2}{d_i d_j}, \tag{3}$$

where m_{ij}'s are the entries of the modularity matrix. Here they are squared, but in the subsequent formula of the Newman–Girvan modularity they are not. The *Newman–Girvan modularity* introduced in [8] directly focuses on modules of higher intra-community connections than expected based on the model of independence.

Definition 1. The Newman–Girvan modularity corresponding to the k-partition $P_k = (V_1, \dots, V_k)$ of the vertex-set of the weighted graph $G = (V, \mathbf{W})$, where the entries of $\mathbf{W}$ sum to 1, is

$$M(P_k, G) = \sum_{a=1}^{k} \sum_{i,j \in V_a} (w_{ij} - d_i d_j).$$

For given integer $1 \leq k \leq n$, the k-module Newman–Girvan modularity of the weighted graph G is

$$M_k(G) = \max_{P_k \in \mathcal{P}_k} M(P_k, G),$$

where $\mathcal{P}_k$ denotes the set of all k-partitions.

For given k, maximizing $M(P_k, G)$ is equivalent to looking for k modules of the vertices with intra-community connections higher than expected under the null-hypothesis. Likewise, minimizing $M(P_k, G)$ is equivalent to looking for k modules of the vertices with intra-community connections lower than expected under the null-hypothesis. It is clear that a 'small' χ^2 value is an indication of a small modularity even in the $k = 1$ case. Therefore, it is worth looking for a modularity structure only after the hypothesis of independence is rejected.

We call a weighted graph *connected* if its vertices cannot be divided into two clusters with all zero between-cluster weights. This is equivalent to the weight matrix being *irreducible* (consequently, all generalized degrees are positive). The modularity matrix $\mathbf{M}$ of (2) always has a zero eigenvalue with eigenvector $\mathbf{1} = \mathbf{1}_n = (1, \ldots, 1)^T$, since its rows sum to zero. Because $\mathrm{tr}(\mathbf{M}) < 0$, $\mathbf{M}$ must have at least one negative eigenvalue, and it is usually indefinite. The *normalized modularity matrix* introduced in [3] is

$$\mathbf{M}_D = \mathbf{D}^{-1/2}\mathbf{M}\mathbf{D}^{-1/2},$$

where $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_n)$ is the diagonal *degree-matrix*. The eigenvalues of $\mathbf{M}_D$ are the same, irrespective of whether we start with the adjacency or normalized edge-weight matrix of an unweighted graph, and they are in the $[-1, 1]$ interval; 1 cannot be an eigenvalue if G is connected.

There are important relations between the eigenvalues of the above matrices, as far as their signs are concerned. The following proposition intensively uses the Sylvester's inertia theorem: if $\mathbf{A}$ is an $n \times n$ symmetric and $\mathbf{B}$ is an $n \times n$ nonsingular matrix, then $\mathbf{B}^T\mathbf{A}\mathbf{B}$ and $\mathbf{A}$ have the same number *pos* of positive, *neg* of negative, and *zero* of zero eigenvalues, i.e., they have the same *inertia* $(pos, neg, zero)$, where $pos + neg + zero = n$.

Proposition 2. *Let $G = (V, \mathbf{W})$ be a connected weighted graph. Its edge-weight matrix $\mathbf{W}$ and normalized edge-weight matrix $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$ have the same inertia $(pos, neg, zero)$ with $pos \geq 1$, whereas its modularity and normalized modularity matrices have inertia $(pos - 1, neg, zero + 1)$.*

Proof. $\mathbf{W}$ and $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$ have the same inertia, since $\mathbf{D}^{-1/2}$ is nonsingular. Likewise, $\mathbf{M}$ and $\mathbf{M}_D$ have the same inertia. Between the inertias of the normalized matrices the following can be established. The spectrum of $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$ is $1 = \mu_0 \geq \mu_1 \geq \cdots \geq \mu_{n-1} \geq -1$ with corresponding unit-norm eigenvectors $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ (see [4]), therefore $pos \geq 1$. Observe that

$$\mathbf{M}_D = \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2} - \sqrt{\mathbf{d}}\sqrt{\mathbf{d}}^T,$$

where $\sqrt{\mathbf{d}} := (\sqrt{d_1}, \dots, \sqrt{d_n})^T$. The eigenvalue 1 of the first term above has multiplicity one with corresponding unit-norm eigenvector $\mathbf{u}_0 = \sqrt{\mathbf{d}}$ whenever $\mathbf{W}$ is irreducible. The only non-zero eigenvalue of the rank 1 second term is also 1 with the same eigenvector. Therefore, if we start with an irreducible $\mathbf{W}$, the spectrum of the matrix $\mathbf{M}_D$ is comprised of $\mu_1 \geq \dots \geq \mu_{n-1} \geq -1$ and the number $\mu_n = 0$ (with eigenvector $\sqrt{\mathbf{d}}$). Consequently, the number of the positive eigenvalues is decreased, while the number of the zero eigenvalues is increased by one. □

The χ^2 statistic of (3) can be written as well in terms of the entries $\tilde{m}_{ij}$'s and the eigenvalues μ_i's of the normalized modularity matrix:

$$\chi^2 = N \sum_{i=1}^{n} \sum_{j=1}^{n} \tilde{m}_{ij}^2 = N \sum_{i=1}^{n} \mu_i^2 = N \sum_{i=1}^{n-1} \mu_i^2,$$

where N is the sample size.

We introduce some further notions. The unweighted graph on n vertices is *complete* if the entries of its adjacency matrix are

$$a_{ij} := \begin{cases} 1 & \text{if } i \neq j \\ 0 & \text{if } i = j. \end{cases}$$

This graph is denoted by K_n.

The unweighted graph on the n-element vertex-set V is *complete multipartite* with $2 \leq k \leq n$ clusters $V_1, \dots, V_k$ (they form a partition of the vertices) if the entries of its adjacency matrix are

$$a_{ij} := \begin{cases} 1 & \text{if } c(i) \neq c(j) \\ 0 & \text{if } c(i) = c(j), \end{cases}$$

where $c(i)$ is the cluster membership of vertex i. Here the non-empty, disjoint vertex-subsets form so-called maximal independent sets of the vertices. If $|V_i| = n_i$ $(i = 1, \dots, k)$, $\sum_{i=1}^{k} n_i = n$, then this graph is denoted by $K_{n_1,\dots,n_k}$.

Note that K_n is also complete multipartite with n singleton clusters, i.e., it is the $K_{1,\dots,1}$ graph. Hence, in the case of $k = n$, the results for complete graphs follow from those for complete multipartites. Therefore, in the sequel, whenever we speak of complete multipartite graphs, complete graphs are also understood, and we always assume that $k \geq 2$. (In the case of $k = 1$, the only cluster would be the empty graph with zero adjacency matrix all zero eigenvalues; further, the notation K_{n_1} would be misleading in this case.)

A weighted graph is called *soft-core* if all its edge-weights are strictly positive (see [6]). Analogously, we will call a weighted graph *soft-core k-partite* with $2 \leq k \leq n$ clusters $V_1, \dots, V_k$ (they form a partition of the vertices) if its edge-weights are

$$w_{ij} = \begin{cases} \text{positive} & \text{if } c(i) \neq c(j) \\ 0 & \text{if } c(i) = c(j), \end{cases}$$

where $c(i)$ is the cluster membership of vertex i. Here the non-empty, disjoint vertex-subsets also form maximal independent sets of the vertices with zero-weighted edges within, and positively weighted edges between them.

In Section 5 we will also use the *normalized Laplacian* of $G = (V, \mathbf{W})$ which is defined as $\mathbf{L}_D = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$. With the notation used in the proof of Proposition 2, the eigenvalues of $\mathbf{L}_D$ are the numbers $\lambda_i = 1 - \mu_i$ $(i = 1, \dots, n-1)$ and $\lambda_0 = 1 - \mu_0 = 0$. Therefore, the spectrum of $\mathbf{L}_D$ is in $[0, 2]$ and 0 is a single eigenvalue if and only if G is connected. The next proposition further characterizes the bottom of the spectrum $0 = \lambda_0 < \lambda_1 \leq \dots \leq \lambda_{n-1} \leq 2$, or equivalently, the top of the spectrum $1 = \mu_0 > \mu_1 \geq \dots \geq \mu_{n-1} \geq -1$ for some special weighted graphs.

Proposition 3. *If the connected weighted graph $G = (V, \mathbf{W})$ has an independent vertex-set of size $1 < k < n$, then its $\mu_{k-1} \geq 0$, or equivalently, $\lambda_{k-1} \leq 1$.*

Proof. Without loss of generality, assume that $w_{ij} = 0$ when $1 \leq i, j \leq k$. Since μ_{k-1} is the kth largest eigenvalue (including the trivial $\mu_0 = 1$) of $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$, the Courant–Fischer–Weyl minimax principle yields that

$$\mu_{k-1} = \max_{\substack{F \subset \mathbb{R}^n \\ \dim(F) = k}} \min_{\substack{\mathbf{x} \in F \\ \|\mathbf{x}\| = 1}} \mathbf{x}^T \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}\mathbf{x}.$$

Therefore, to prove that $\mu_{k-1} \geq 0$, it suffices to find a k-dimensional subspace $F \subset \mathbb{R}^n$ such that $\min_{\substack{\mathbf{x} \in F \\ \|\mathbf{x}\| = 1}} \mathbf{x}^T \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}\mathbf{x} = 0$. Set $F := \{\mathbf{x} : \mathbf{x} = (x_1, \dots, x_k, 0, \dots, 0) \in \mathbb{R}^n\}$. Clearly, for every $\mathbf{x} \in F$: $\mathbf{x}^T \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}\mathbf{x} = 0$, and this also holds true for unit-norm $\mathbf{x}$'s. Therefore, the above minimum is also 0. This, together with the relation $\lambda_{k-1} = 1 - \mu_{k-1}$, finishes the proof. □

By Proposition 3, the case $k = 2$ implies that $\mu_1 \geq 0$, or equivalently, $\lambda_1 \leq 1$ whenever G is not a soft-core weighted graph, i.e., it has at least one 0 weight. On the other hand, $\mu_1 < 0$ (and so $\lambda_1 > 1$) when there are not two independent vertices in the graph; particularly, among unweighted graphs, in the case of K_n (see the forthcoming Proposition 5).

Actually, in the case of $K_{n_1,\dots,n_k}$ $(k < n)$, the nonnegative μ_i's, guaranteed by Proposition 3, are all zeros, or equivalently, the smallest positive eigenvalue of its normalized Laplacian is 1 with multiplicity $n - k$, see Table 3.1 of [4]. In Section 4 we will prove that, among unweighted graphs, equality is attained ($\mu_1 = 0$, $\lambda_1 = 1$) only for $K_{n_1,\dots,n_k}$ $(k < n)$. We will also prove that whenever a weighted graph is not soft-core multipartite, it must have $\mu_1 > 0$, or equivalently, $\lambda_1 < 1$.

Since the investigated matrices are closely related, the statements of some theorems of Sections 3 and 4 occasionally follow from others, and vice versa. However, if the proofs contain different ideas, we include them both and allow the reader to decide which is of greater concern to him or her. In this way we intend to give a deeper insight into the investigated graphs and the related matrices.

3. Modularity spectra of complete and complete multipartite graphs

Now we derive the modularity spectra of the graphs introduced in the previous section.

Proposition 4. *The spectrum of $\mathbf{M}(K_n)$ consists of the single eigenvalue 0 with eigenvector $\mathbf{1}_n$ and the number -1 with multiplicity $n-1$ and eigen-subspace $\mathbf{1}_n^{\perp}$.*

Proof. The adjacency matrix of K_n is $\mathbf{A}(K_n) = \mathbf{1}_n\mathbf{1}_n^T - \mathbf{I}_n$; further, $\mathbf{d} = (n-1)\mathbf{1}_n$ and $2e = n(n-1)$. Hence, in view of (1),

$$\begin{aligned}
\mathbf{M}(K_n) &= \mathbf{1}_n\mathbf{1}_n^T - \mathbf{I}_n - \frac{n-1}{n}\mathbf{1}_n\mathbf{1}_n^T = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T - \mathbf{I}_n \\
&= \left(\frac{\mathbf{1}}{\sqrt{n}}\right)\left(\frac{\mathbf{1}}{\sqrt{n}}\right)^T - \left[\left(\frac{\mathbf{1}}{\sqrt{n}}\right)\left(\frac{\mathbf{1}}{\sqrt{n}}\right)^T + \sum_{i=2}^{n} 1 \cdot \mathbf{u}_i\mathbf{u}_i^T\right] \\
&= \sum_{i=2}^{n}(-1)\cdot \mathbf{u}_i\mathbf{u}_i^T,
\end{aligned}$$

where $\mathbf{u}_2, \dots, \mathbf{u}_n$ is an arbitrary orthonormal set in $\mathbf{1}_n^{\perp}$. Therefore, the unique spectral decomposition of $\mathbf{M}(K_n)$ is as stated in the proposition. □

Proposition 5. *The spectrum of $\mathbf{M}_D(K_n)$ consists of the single eigenvalue 0 with eigenvector $\mathbf{1}_n$ and the number $-\frac{1}{n-1}$ with multiplicity $n-1$ and eigen-subspace $\mathbf{1}_n^{\perp}$.*

This proposition follows from the fact that $\mathbf{M}_D(K_n) = \frac{1}{n-1}\mathbf{M}(K_n)$.

Proposition 6. *The modularity spectrum of the complete multipartite graph $K_{n_1,\dots,n_k}$ consists of $k-1$ strictly negative eigenvalues and zero with multiplicity $n-k+1$.*

Proof. The adjacency matrix of $K_{n_1,\dots,n_k}$ is a block-matrix with diagonal blocks of all zeros and off-diagonal blocks of all 1's. Let $V_1, \dots, V_k$ denote the independent, disjoint vertex-subsets (clusters), $|V_i| = n_i$, $i = 1, \dots, k$; $d_j = n - n_i$ if $j \in V_i$; $2e = \sum_{j=1}^{n} d_j = \sum_{i=1}^{k} n_i(n-n_i) = n^2 - \sum_{i=1}^{k} n_i^2$. Therefore, $\mathbf{M} = \mathbf{M}(K_{n_1,\dots,n_k})$ is also a block-matrix, where the entries in the block of size $n_i \times n_j$ are all equal to the following number p_{ij}:

$$p_{ij} = (1-\delta_{ij}) - \frac{(n-n_i)(n-n_j)}{2e}, \quad i,j = 1,\dots,k.$$

Here δ_{ij} stands for the Kronecker delta-symbol. With the wording of [4], $\mathbf{M}$ is a *blown-up matrix* with blow-up sizes $n_1, \dots, n_k$ of the $k \times k$ symmetric pattern matrix $\mathbf{P}$ of entries p_{ij}'s. Consequently, $\operatorname{rank}(\mathbf{M}) = \operatorname{rank}(\mathbf{P}) \le k$. We will prove that $\mathbf{P}$ has rank exactly $k-1$, and all its nonzero eigenvalues are strictly negative.

Let us look for an eigenvector $\mathbf{u}$ of $\mathbf{M}$, belonging to a nonzero eigenvalue λ, of piecewise constant coordinates over the partition corresponding to $V_1, \dots, V_k$ of $\{1, \dots, n\}$, i.e., let n_i coordinates be equal to y_i, $i = 1, \dots, k$. With these, the eigenvalue–eigenvector equation yields that

$$\sum_{j=1}^{k} p_{ij} n_j y_j = \lambda y_i. \tag{4}$$

Therefore, λ is an eigenvalue of the $k \times k$ matrix $\mathbf{PN}$ with eigenvector $(y_1, \dots, y_k)^T$, where $\mathbf{N} = \operatorname{diag}(n_1, \dots, n_k)$. The matrix $\mathbf{PN}$ is not symmetric, but its eigenvalues are real because they are originally eigenvalues of an $n \times n$ symmetric matrix, or else, its eigenvalues are also eigenvalues of the $k \times k$ symmetric matrix $\mathbf{N}^{1/2}\mathbf{P}\mathbf{N}^{1/2}$. It is easy to see that the row sums of $\mathbf{PN}$ are zeros, since

$$\sum_{j=1}^{k} p_{ij} n_j = \sum_{j=1}^{k} \left[(1 - \delta_{ij}) - \frac{(n - n_i)(n - n_j)}{2e} \right] n_j = 0. \tag{5}$$

Therefore, zero is an eigenvalue of $\mathbf{PN}$ with eigenvector $\mathbf{1}_k$, which results in another zero eigenvalue of $\mathbf{M}$ with eigenvector $\mathbf{1}_n$. Thus, zero is an eigenvalue of $\mathbf{M}$ with multiplicity at least $n - k + 1$.

Now we will prove that all the nonzero eigenvalues of $\mathbf{PN}$ are negative. Let $\lambda \neq 0$ be an eigenvalue of $\mathbf{PN}$. In view of (4) and (5),

$$\lambda \sum_{i=1}^{k} n_i y_i = \sum_{i=1}^{k} n_i (\lambda y_i) = \sum_{j=1}^{k} n_j y_j \sum_{i=1}^{k} n_i p_{ij} = 0.$$

Consequently, if $\lambda \neq 0$, then $\sum_{i=1}^{k} n_i y_i = 0$. Now consider

$$\lambda \sum_{i=1}^{k} n_i y_i^2 = \sum_{i=1}^{k} (n_i y_i)(\lambda y_i) = \sum_{i=1}^{k} n_i y_i \sum_{j=1}^{k} p_{ij} n_j y_j$$
$$= \sum_{i=1}^{k} \sum_{j=1}^{k} p_{ij} (n_i y_i)(n_j y_j).$$

We will show that the right hand side is negative, and therefore, by $\sum_{i=1}^{k} n_i y_i^2 > 0$, we get that $\lambda < 0$. Indeed,

$$\begin{aligned}\sum_{i=1}^{k}\sum_{j=1}^{k} p_{ij}(n_i y_i)(n_j y_j) &= \sum_{i=1}^{k}\sum_{j=1}^{k}\left[(1-\delta_{ij}) - \frac{(n-n_i)(n-n_j)}{2e}\right](n_i y_i)(n_j y_j) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k}(1-\delta_{ij})(n_i y_i)(n_j y_j) \\ &\quad - \frac{1}{2e}\left[\sum_{i=1}^{k}(n-n_i)n_i y_i\right]\left[\sum_{j=1}^{k}(n-n_j)n_j y_j\right] \\ &= \left(\sum_{i=1}^{k} n_i y_i\right)\left(\sum_{j=1}^{k} n_j y_j\right) - \sum_{i=1}^{k}(n_i y_i)^2 \\ &\quad - \frac{1}{2e}\left[\sum_{i=1}^{k}(n-n_i)n_i y_i\right]^2 < 0,\end{aligned}$$

where we used that $\sum_{i=1}^{k} n_i y_i = 0$.

Since the piecewise constant vectors (over $V_1, \dots, V_k$), which are also orthogonal to the $\mathbf{1}_n$ vector, constitute a $(k-1)$-dimensional subspace of $\mathbb{R}^n$, there should be $k-1$ strictly negative eigenvalues of $K_{n_1,\dots,n_k}$. Consequently, the eigenvalue zero of $\mathbf{M}$ has multiplicity exactly $n-k+1$ with corresponding eigen-subspace

$$\left\{\mathbf{x} : \sum_{j\in V_i} x_j = 0,\, i = 1,\dots,k\right\},$$

which includes the vector $\mathbf{1}_n$ as well. Thus, we proved that $\mathbf{M}(K_{n_1,\dots,n_k})$ is negative semidefinite. $\square$

Remark 7. In the proof of Proposition 6 we described the eigenvectors and eigen-subspaces of $\mathbf{M} = \mathbf{M}(K_{n_1,\dots,n_k})$ too. The inertia of $\mathbf{M}$ itself can be concluded with the following simple argument, suggested by the anonymous referee of this paper. Observe that the adjacency matrix of $K_{n_1,\dots,n_k}$, with independent vertex-sets $V_1, \dots, V_k$ and $\sum_{i=1}^{k} n_i = n$, can be decomposed as

$$\mathbf{A}(K_{n_1,\dots,n_k}) = (\mathbf{1}_{V_1}, \dots, \mathbf{1}_{V_k}) \cdot \mathbf{A}(K_k) \cdot (\mathbf{1}_{V_1}, \dots, \mathbf{1}_{V_k})^T,$$

where $\mathbf{1}_{V_i} \in \mathbb{R}^n$ is the indicator vector of V_i ($i = 1, \dots, k$), and $\mathbf{A}(K_k)$ is the adjacency matrix of the complete graph on k vertices (with the wording of the previous proof, we blow it up). The vectors $\mathbf{b}_1, \dots, \mathbf{b}_{n-k}$ are chosen so that together with the indicator vectors they form a complete orthogonal basis in $\mathbb{R}^n$. Then, with $\mathbf{B} = (\mathbf{1}_{V_1}, \dots, \mathbf{1}_{V_k}, \mathbf{b}_1, \dots, \mathbf{b}_{n-k})$, the matrix $\mathbf{A}$ is congruent to the following:

$$\begin{pmatrix} \mathbf{A}(K_k) & \mathbf{O}_{k\times(n-k)} \\ \mathbf{O}_{(n-k)\times k} & \mathbf{O}_{k\times k} \end{pmatrix}, \tag{6}$$

where $\mathbf{O}$ denotes the matrix of all zeros, and the inertia of $\mathbf{A}(K_k) = \mathbf{1}_k\mathbf{1}_k^T - \mathbf{I}_k$ is $(1, k-1, 0)$, which can easily be seen from the proof of Proposition 4. Consequently, the inertia of the matrix in (6) and that of the adjacency matrix of $K_{n_1,\dots,n_k}$ is $(1, k-1, n-k)$. In view of Proposition 2, the inertia of the modularity matrix of $K_{n_1,\dots,n_k}$ is therefore $(0, k-1, n-k+1)$.

Proposition 8. $\mathbf{M}_D(K_{n_1,\dots,n_k})$ *is also negative semidefinite.*

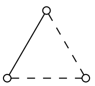
The proof follows by Proposition 2.

4. The main statements

To prove our main statements, we will extensively use the following well-known characterization of the complete multipartite graphs (including the complete graphs): an unweighted connected graph is complete multipartite if and only if it has no three-vertex induced subgraph with exactly one edge. More generally, we are able to give a similar characterization for weighted soft-core multipartite graphs.

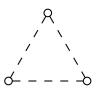
Lemma 9. *A weighted graph is soft-core multipartite if and only if it has no triangle with exactly one positively weighted edge.*

Proof. We will call the above triangle *forbidden pattern*, which looks like

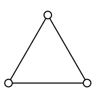


(the solid line means an edge of positive weight, whereas the dashed one means an edge of zero weight).

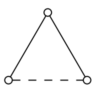
- In the forward direction, a soft-core multipartite graph can have the following types of triangles (not all of them appear necessarily, only if the size of clusters allows it):
  (1) the three vertices are from the same cluster, in which case the triangle has edges of zero weight;

  (2) the three vertices are from three different clusters, in which case the triangle is a soft-core graph of all positive edges;

  (3) two of the vertices are from the same, and the third from a different cluster, in which case the triangle has exactly two edges of positive weight (cherry).

  None of them is the forbidden pattern.

- Conversely, suppose that our weighted graph does not have the forbidden pattern. The following procedure shows that it is then soft-core multipartite. Let the first cluster be a maximal independent set of the vertices, say V_1. We claim that each vertex in $\overline{V}_1$ is connected with a positively weighted edge to each vertex of V_1. Indeed, let $c \in \overline{V}_1$ be a vertex; there must be a vertex (say, a) of V_1 with $w_{ca} > 0$, since if not, it could be joined to V_1, which contradicts the maximality of V_1 as an independent set. If c were not connected to another $b \in V_1$ with a positively weighted edge, then a, b, c would form a forbidden pattern, but our graph does not contain such in view of our starting assumption.
  Then let V_2 be a maximal independent set of vertices within $\overline{V}_1$, say V_2. We claim that each vertex in $\overline{V_1 \cup V_2}$ is connected to each vertex of V_1 and V_2 with a positively weighted edge. The connectedness to vertices of V_1 is already settled. By the maximality of V_2 as an independent set, any vertex of $\overline{V_1 \cup V_2}$ must be connected to at least one vertex of V_2 with a positively weighted edge. If we found a vertex $c \in \overline{V_1 \cup V_2}$ such that for some $a \in V_2$: $w_{ac} > 0$, and for another $b \in V_2$: $w_{bc} = 0$, then a, b, c would form a forbidden pattern, which is excluded.
  Advancing in this way, one can see that the procedure produces maximal disjoint independent sets of the vertices such that the independent vertices of V_k are connected to every vertex in $V_1, \dots, V_{k-1}$. At each step we can select a maximal independent set out of the remaining vertices; in the worst case it contains only one vertex. The absence of the forbidden pattern guarantees that we can always continue our algorithm until all vertices are placed into a cluster. This procedure will exhaust the set of vertices and result in a soft-core multipartite graph. The point is that in the absence of the forbidden pattern we can divide the vertices into independent sets which are fully connected. $\square$

Note that if we proceed with non-increasing cardinalities of V_i's, then one-vertex independent sets may emerge at the end of the process. Moreover, up to the labeling of the vertices and the numbering of the independent sets, the resulting soft-core multipartite structure is unique. In fact, the above procedure just recovers this unique structure in the absence of the forbidden pattern.

Now, we are able to prove the following.

Theorem 10. *If the connected weighted graph $G = (V, \mathbf{W})$ is not soft-core multipartite, then the largest eigenvalue of its modularity matrix is strictly positive.*

Proof. By Lemma 9, a weighted graph is not soft-core multipartite if and only if it contains the forbidden pattern. Let us consider such a graph. Since the modularity spectrum does not depend on the labeling of the vertices, assume that the first three vertices form the forbidden pattern, i.e., the upper left corner of the edge-weight matrix is

$$\begin{pmatrix} 0 & w_{12} & 0 \\ w_{21} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

with $w_{12} = w_{21} > 0$. Based on (2), this graph's modularity matrix is

$$\mathbf{M} = \frac{1}{c^2}(c\mathbf{W} - \mathbf{d}\mathbf{d}^T),$$

where $\mathbf{d} = (d_1, \dots, d_n)^T$, $d_i = \sum_{j=1}^n w_{ij}$ $(i = 1, \dots, n)$, and $c = \sum_{i=1}^n d_i$ (not necessarily an integer).

It is known that a matrix is negative semidefinite if and only if its every principal minor of odd order is non-positive, and every principal minor of even order is non-negative. The principal minor of order 3 of $\mathbf{M}$ is

$$\begin{aligned} &\left(\frac{1}{c^2}\right)^3 \det \begin{pmatrix} -d_1^2 & cw_{12} - d_1 d_2 & -d_1 d_3 \\ cw_{12} - d_1 d_2 & -d_2^2 & -d_2 d_3 \\ -d_1 d_3 & -d_2 d_3 & -d_3^2 \end{pmatrix} \\ &= \frac{1}{c^6} c^2 w_{12}^2 d_3^2 = \frac{w_{12}^2 d_3^2}{c^4} > 0. \end{aligned}$$

Because G is connected, both d_3 and c are strictly positive, akin to the above odd-order minor. Consequently, the modularity matrix cannot be negative semidefinite, hence it must contain at least one positive eigenvalue. □

Theorem 10 together with Proposition 6 gives a necessary and sufficient condition for an unweighted graph to have the zero as the largest eigenvalue of its modularity matrix.

Theorem 11. *The modularity matrix of an unweighted connected graph is negative semidefinite if and only if it is complete multipartite.*

By Proposition 2, the same statement holds for the normalized modularity matrix. Although it follows from Theorem 10, we will give an alternative proof of the forthcoming Theorem 12, since it contains fewer calculations and may be more alluring for the reader. The reader may also note that Theorem 12 could be stated first, in which case Theorem 10 is implied by Theorem 12.

Theorem 12. *If the connected weighted graph $G = (V, \mathbf{W})$ is not soft-core multipartite, then the largest eigenvalue of its normalized modularity matrix is strictly positive.*

Proof. Referring to Section 2, the largest eigenvalue μ_1 of $\mathbf{M}_D$ is the second largest eigenvalue of $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$, whose largest eigenvalue is 1 with corresponding eigenvector $\sqrt{\mathbf{d}}$ (this is unique if our graph is connected). Therefore, we think in terms of the two largest eigenvalues of $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$. We can again assume that the first three vertices form the forbidden pattern and so, the upper left corner of this matrix looks like

$$\begin{pmatrix} 0 & \frac{w_{12}}{\sqrt{d_1 d_2}} & 0 \\ \frac{w_{21}}{\sqrt{d_1 d_2}} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

with $w_{12} = w_{21} > 0$.

Then the Courant–Fischer–Weyl minimax principle yields

$$\mu_1 = \max_{\substack{\|\mathbf{x}\|=1 \\ \mathbf{x}^T\sqrt{\mathbf{d}}=0}} \mathbf{x}^T \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}\mathbf{x}.$$

Therefore, to prove that $\mu_1 > 0$, it suffices to find an $\mathbf{x} \in \mathbb{R}^n$ that satisfies conditions $\|\mathbf{x}\| = 1$, $\mathbf{x}^T\sqrt{\mathbf{d}} = 0$, and for which, $\mathbf{x}^T\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}\mathbf{x} > 0$. (The unit norm condition can be relaxed here, because $\mathbf{x}$ can later be normalized, without changing the sign of the above quadratic form.)

Indeed, let us look for $\mathbf{x}$ of the form $\mathbf{x} = (x_1, x_2, x_3, 0, \dots, 0)^T$ such that

$$\sqrt{d_1}x_1 + \sqrt{d_2}x_2 + \sqrt{d_3}x_3 = 0. \tag{7}$$

Then the inequality

$$\mathbf{x}^T\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}\mathbf{x} = \frac{2x_1x_2w_{12}}{\sqrt{d_1d_2}} > 0$$

can be satisfied with any $\mathbf{x} = (x_1, x_2, x_3, 0, \dots, 0)^T$ such that x_1 and x_2 are both positive or both negative, in which case, due to (7),

$$x_3 = -\frac{\sqrt{d_1}x_1 + \sqrt{d_2}x_2}{\sqrt{d_3}}$$

is a good choice, and will have the opposite sign. (Note that all the d_i's are positive, since we deal with connected weighted graphs.) □

Theorem 12 together with Proposition 8 gives the following statement of equivalence.

Theorem 13. *The normalized modularity matrix of an unweighted connected graph is negative semidefinite if and only if it is complete multipartite.*

5. Applications

The results of Theorems 10, 12, and 13 have important implications in the following, seemingly unrelated areas.

5.1. The Newman–Girvan modularity

Based on Definition 1, the two-way Newman–Girvan modularity of $G = (V, \mathbf{W})$ is

$$M_2(G) = \max_{\substack{U \subset V \\ U \neq \emptyset,\, U \neq V}} M_2\big((U, \overline{U}), G\big),$$

where the modularity of the proper two-partition $(U, \overline{U})$ of V can be written in terms of the entries m_{ij}'s (summing to 0) of the modularity matrix of G as follows:

$$\begin{aligned} M_2\big((U, \overline{U}), G\big) &= \sum_{i,j \in U} m_{ij} + \sum_{i,j \in \overline{U}} m_{ij} = -2 \sum_{i \in U,\, j \in \overline{U}} m_{ij} \\ &= -2\big[w(U, \overline{U}) - \mathrm{Vol}(U)\,\mathrm{Vol}(\overline{U})\big], \end{aligned} \tag{8}$$

where $w(U, \overline{U}) = \sum_{i \in U} \sum_{j \in \overline{U}} w_{ij}$ is the *weighted cut* between U and $\overline{U}$, and $\mathrm{Vol}(U) = \sum_{i \in U} d_i$ is the *volume* of the vertex-subset U. These formulas are valid under the condition $\mathrm{Vol}(V) = 1$.

Now we use the idea of the proof of the *Expander Mixing Lemma* (see [7]) extended to weighted graphs (see [5]).

Lemma 14. *Let $G = (V, \mathbf{W})$ be a weighted graph with $\mathrm{Vol}(V) = 1$. Then for all $X, Y \subset V$:*

$$\big|w(X, Y) - \mathrm{Vol}(X)\,\mathrm{Vol}(Y)\big| \leq \|\mathbf{M}_D\| \cdot \sqrt{\mathrm{Vol}(X)\,\mathrm{Vol}(Y)},$$

where $\|\mathbf{M}_D\|$ is the spectral norm (the largest absolute value of the eigenvalues) of the normalized modularity matrix of G.

With the notation of the proof of this lemma (see [5]), and introducing $\mu_0 = 1$, $\mathbf{u}_0 = \sqrt{\mathbf{d}}$,

$$\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2} = \sum_{i=0}^{n-1} \mu_i \mathbf{u}_i \mathbf{u}_i^T$$

is a spectral decomposition.

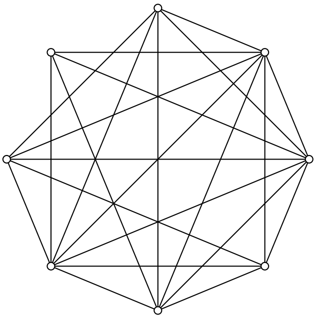
Let $U \subset V$ be arbitrary and the indicator vector of U is denoted by $\mathbf{1}_U \in \mathbb{R}^n$. Further, set $\mathbf{x} := \mathbf{D}^{1/2}\mathbf{1}_U$ and $\mathbf{y} := \mathbf{D}^{1/2}\mathbf{1}_{\overline{U}}$, and let $\mathbf{x} = \sum_{i=0}^{n-1} a_i \mathbf{u}_i$ and $\mathbf{y} = \sum_{i=0}^{n-1} b_i \mathbf{u}_i$ be the expansions of $\mathbf{x}$ and $\mathbf{y}$ in the orthonormal basis $\mathbf{u}_0, \dots, \mathbf{u}_{n-1}$ with coordinates $a_i = \mathbf{x}^T \mathbf{u}_i$ and $b_i = \mathbf{y}^T \mathbf{u}_i$, respectively. Observe that $w(U, \overline{U}) = \mathbf{1}_U^T \mathbf{W} \mathbf{1}_{\overline{U}} = \mathbf{x}^T(\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2})\mathbf{y}^T$ and $\mathbf{1}_{\overline{U}} = \mathbf{1}_n - \mathbf{1}_U$; therefore,

$$\begin{aligned} b_i &= \mathbf{y}^T \mathbf{u}_i = \mathbf{D}^{1/2}(\mathbf{1} - \mathbf{1}_U)\mathbf{u}_i \\ &= \mathbf{u}_0^T \mathbf{u}_i - \mathbf{x}^T \mathbf{u}_i = -a_i \quad (i = 1, 2, \dots, n-1); \end{aligned}$$

further, $a_0 = \mathrm{Vol}(U)$ and $b_0 = \mathrm{Vol}(\overline{U})$. Based on these observations,

$$w(U, \overline{U}) - \mathrm{Vol}(U)\,\mathrm{Vol}(\overline{U}) = \sum_{i=1}^{n-1} \mu_i a_i b_i = -\sum_{i=1}^{n-1} \mu_i a_i^2.$$

Consequently, by (8), $M_2((U, \overline{U}), G) = 2\sum_{i=1}^{n-1} \mu_i a_i^2$. Therefore, provided that the normalized modularity matrix of the underlying weighted graph is negative semidefinite (or equivalently, our graph is complete multipartite when it is unweighted), $M_2((U, \overline{U}), G) \le 0$ for all two-partitions of the vertices, and hence, the two-way Newman–Girvan modularity, $M_2(G)$, is also non-positive (in most cases, it is negative). Nonetheless, this property does not characterize the complete multipartite graphs. There are graphs with positive μ_1 that have zero or sometimes negative two-way Newman–Girvan modularity. Therefore, the negative semidefiniteness of the modularity matrix is not necessary for a graph to be indivisible. For example, consider the following unweighted graph obtained by deleting 5 edges from the complete graph on 8 vertices:



The largest eigenvalue of the corresponding modularity matrix is positive (0.6725), but the corresponding maximum two-way modularity is negative: $M_2(G) = -0.0076$; therefore, our graph is indivisible. Note that here the spectrum is shifted to the negative direction (the smallest eigenvalue is -2.3324, and there are four additional negative eigenvalues, while the zero eigenvalue has multiplicity two).

In [4] we discuss how the balance of the negative and positive eigenvalues with large absolute value determines a community, anti-community, or just a regular structure.

5.2. The isoperimetric number

Due to Theorem 13 and the relation between the normalized Laplacian and modularity spectra, the smallest positive normalized Laplacian eigenvalue, λ_1, is slightly greater than 1 for complete, equal to 1 for complete multipartite (but not complete), and strictly less than 1 for other unweighted graphs. In the weighted case, in view of Theorem 12, $\lambda_1 < 1$ whenever our graph is not soft-core multipartite. In [2] we gave the following upper and lower estimate for the Cheeger constant $h(G) = \min_{\mathrm{Vol}(U) \le \frac{1}{2}} \frac{w(U, \overline{U})}{\mathrm{Vol}(U)}$ of the weighted graph $G = (V, \mathbf{W})$ by its smallest positive normalized Laplacian eigenvalue:

$$\frac{\lambda_1}{2} \le h(G) \le \sqrt{\lambda_1(2-\lambda_1)}$$

whenever $\lambda_1 \le 1$. Therefore, if λ_1 is separated from zero but less than 1, then it is an indication of the high edge-expansion of a not soft-core multipartite graph. The upper estimate is not valid for complete graphs (for which the lower bound is attained), and it gives the trivial upper bound 1 for complete bipartite or multipartite (but not complete) graphs. Indeed, the former are, in fact, super-expanders, while the latter are so-called bipartite or multipartite expanders, see [1]. For large n, the situation can be even more complicated and also influenced by the large absolute value negative eigenvalues of the normalized modularity matrix. More generally, one may look for so-called volume-regular cluster pairs of small discrepancy by means of the eigenvectors corresponding to the structural (large absolute value) eigenvalues of $\mathbf{M}_D$ (see [5]).

5.3. The symmetric maximal correlation

The largest eigenvalue of $\mathbf{M}_D$, μ_1, is called *symmetric maximal correlation* in the setting of correspondence analysis on the symmetric contingency table $\mathbf{W}$ (see [2]). Indeed, the weight matrix $\mathbf{W}$ (with sum of its entries 1) defines a symmetric discrete joint distribution $\mathbb{W}$ with equal margins $\mathbb{D} = \{d_1, \ldots, d_n\}$. Let H denote the Hilbert space of $V \to \mathbb{R}$ random variables taking on at most n different values with probabilities $d_1, \ldots, d_n$, further, having zero expectation. Let us take two identically distributed (i.d.) copies $\psi, \psi' \in H$ with joint distribution $\mathbb{W}$. The symmetric maximal correlation with respect to $\mathbb{W}$ is the following:

$$r_1 = \max_{\psi,\psi' \in H \text{ i.d.}} \mathrm{Corr}_{\mathbb{W}}(\psi, \psi') = \max_{\substack{\psi,\psi' \in H \text{ i.d.} \\ \mathrm{Var}_{\mathbb{D}}\psi = 1}} \mathrm{Cov}_{\mathbb{W}}(\psi, \psi').$$

In [2] we proved that $r_1 = 1 - \lambda_1$, i.e., $r_1 = \mu_1$.

Then Theorem 12 implies that r_1 is strictly positive if the joint distribution is not of a soft-core multipartite structure, i.e., it contains the forbidden pattern. Note that the existence of such a pattern is not a particular requirement. In the setting of social networks, it can be interpreted in the following way: there is a set of three people such that two among them are connected, while the third is not connected to either of the other two people.

In the case of a binary table, the converse is also true, and hence, the symmetric maximal correlation is positive if and only if the joint distribution is not of a complete multipartite structure.

Acknowledgements

This article was written under the auspices of the Budapest Semesters of Mathematics program, in the framework of an undergraduate research course on spectral clustering with the participation of US students: Brian Bullins from Duke University, Durham, NC; Sorathan Chaturapruek from Harvey Mudd College, Claremont, CA; and Shiwen Chen from Williams College, Williamstown, MA. Marianna Bolla's research was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project, sponsored by the European Union. Katalin Friedl's research was partially supported by the Hungarian National Research Grant OTKA-108947.

We thank the anonymous referee of this paper for many valuable remarks and recommendations; particularly, Remark 7 and the current stronger version of Proposition 3 are his/her contributions.

References

[1] N. Alon, Eigenvalues and expanders, Combinatorica 6 (1986) 83–96.
[2] M. Bolla, G. Molnár-Sáska, Optimization problems for weighted graphs and related correlation estimates, Discrete Math. 282 (2004) 23–33.
[3] M. Bolla, Penalized versions of the Newman–Girvan modularity and their relation to multiway cuts and k-means clustering, Phys. Rev. E 84 (2011) 016108.
[4] M. Bolla, Spectral Clustering and Biclustering. Learning Large Graphs and Contingency Tables, Wiley, 2013.
[5] M. Bolla, Modularity spectra, eigen-subspaces and structure of weighted graphs, European J. Combin. 35 (2014) 105–116.
[6] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi, Convergent graph sequences I: subgraph frequencies, metric properties, and testing, Adv. Math. 219 (2008) 1801–1851.
[7] S. Hoory, N. Linial, A. Widgerson, Expander graphs and their applications, Bull. Amer. Math. Soc. (N.S.) 43 (4) (2006) 439–561.
[8] M.E.J. Newman, M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2004) 026113.
[9] M.E.J. Newman, Modularity and community structure in networks, Proc. Natl. Acad. Sci. USA 103 (2006) 8577–8582.
[10] M.E.J. Newman, Finding community structure in networks using the eigenvectors of matrices, Phys. Rev. E 74 (2006) 036104.