Journal of
Mathematical
Inequalities
Volume 11, Number 3 (2017), 701–714 doi:10.7153/jmi-2017-11-56

MODULARITY BOUNDS FOR CLUSTERS LOCATED BY LEADING EIGENVECTORS OF THE NORMALIZED MODULARITY MATRIX

DARIO FASINO AND FRANCESCO TUDISCO

(*Communicated by J. Pečarić*)

Abstract. Nodal theorems for generalized modularity matrices ensure that the cluster located by the positive entries of the leading eigenvector of various modularity matrices induces a connected subgraph. In this paper we obtain lower bounds for the modularity of that subgraph showing that, under certain conditions, the nodal domains induced by eigenvectors corresponding to highly positive eigenvalues of the normalized modularity matrix have indeed positive modularity, that is, they can be recognized as modules inside the network. Moreover we establish Cheeger-type inequalities for the cut-modularity of the graph, providing a theoretical support to the common understanding that highly positive eigenvalues of modularity matrices are related with the possibility of subdividing a network into communities.

1. Introduction

The study of community structures in complex networks is facing a significant growth, as observations on real life graphs reveal that many social, biological, and technological networks are intrinsically divided into clusters. Given a generic graph describing some kind of relationship among actors of a complex network, community detection problems basically consist in discovering and revealing the groups (if any) in which the network is subdivided.

Modularity matrices, the main subject of investigation of the present work, are a relevant tool in the development of a sound theoretical background of community detection. Even though a number of modularity matrices has been proposed so far, see e.g., [9] and the references therein, the original and most popular one was introduced by Newman and Girvan in [17] and is defined as a particular rank-one modification of the adjacency matrix. We shall refer to such matrix as the *Newman–Girvan* (or *unnormalized*) modularity matrix, and we will introduce consequently a normalized version of that matrix.

Spectral algorithms are widely applied to data clustering problems, including finding communities or partitions in graphs and networks. In the latter case, sign patterns in the entries of certain eigenvectors of Laplacian matrices are exploited to build vertex subsets, called nodal domains, which often yield excellent solutions to certain combinatorial problems related to the optimal partitioning of a given graph.

Mathematics subject classification (2010): 05C50, 15A18, 15B99.
Keywords and phrases: Nodal domain, community detection, modularity, Cheeger inequality.

Paper JMI-11-56

Analogously, nodal domains of modularity matrices play a crucial role in the community detection framework. A nodal domain theorem has been proved for these matrices [8, 9] showing the connectedness properties of nodal domains associated with their eigenvectors. The main results of this paper show that, under certain conditions, the nodal domains induced by eigenvectors corresponding to positive eigenvalues of the *normalized* modularity matrix have indeed positive modularity, that is, they can be recognized as modules inside the graph. Moreover, we prove two Cheeger-type inequalities for the cut-modularity providing a theoretical support to the common understanding that highly positive eigenvalues of modularity matrices are related with the possibility of subdividing the graph into communities.

The paper is organized as follows. After fixing our notation and preliminary results, in Section 2 we introduce with more detail the modularity based community detection problem, motivating our subsequent investigations. In Section 3 we discuss the unnormalized and normalized versions of the Newman–Girvan modularity matrix, summarizing some of their main structural properties, and we present our main results, concerning the relation between positive eigenvalues of the normalized modularity matrix and modules inside the graph. In Section 4 we prove two Cheeger-type inequalities for the cut-modularity of the graph. Section 5 contains complementary results on modularity properties of nodal domains corresponding to positive eigenvalues of the normalized modularity matrix.

1.1. Notations and preliminaries

In the sequel we give a brief review of standard concepts and symbols from algebraic graph theory that we will use throughout the paper. We assume that $G = (V,E)$ is a finite, undirected, connected, unweighted graph without multiple edges, where V and E are the vertex and edge sets, respectively. We will identify V with $\{1,\dots,n\}$. We denote adjacency of vertices x and y as $xy \in E$. For any $i \in V$, let d_i denote its degree. Moreover, we let $d = (d_1,\dots,d_n)^T$, and $D = \mathrm{Diag}(d_1,\dots,d_n)$. The *average degree* is $\langle d\rangle = (\sum_{i=1}^n d_i)/n$.

The symbols A and $\mathscr{A}$ denote the adjacency matrix of G and its normalized counterpart, that is, $A = (a_{ij})$ where $a_{ij} = 1$ if $ij \in E$, and $a_{ij} = 0$ otherwise; and $\mathscr{A} = D^{-1/2}AD^{-1/2}$. In particular, both A and $\mathscr{A}$ are symmetric, irreducible, componentwise nonnegative matrices. The spectral radius of A is denoted by $\rho(A)$, and $\mathbb{1}$ denotes an all-one vector whose dimension depends on the context.

The cardinality of a set S is denoted by $|S|$. In particular, $|V| = n$. For any $S \subseteq \{1,\dots,n\}$ let $\mathbb{1}_S$ be its *characteristic vector*, defined as $(\mathbb{1}_S)_i = 1$ if $i \in S$ and $(\mathbb{1}_S)_i = 0$ otherwise. Moreover, we denote by $\bar{S}$ the complement $V \setminus S$, and let $\mathrm{vol}\, S = \sum_{i\in S} d_i$ be the *volume* of S. Correspondingly, $\mathrm{vol}\, V = \sum_{i\in V} d_i$ denotes the volume of the whole graph. For any subsets $S,T \subseteq V$ let

$$e(S,T) = \mathbb{1}_S^T A \mathbb{1}_T.$$

For simplicity, we use the shorthands $e_{\mathrm{in}}(S) = e(S,S)$ and $e_{\mathrm{out}}(S) = e(S,\bar{S})$, so that $e_{\mathrm{in}}(S)$ is twice the number of inner-edges in S and $e_{\mathrm{out}}(S)$ is the size of the edgeboundary of S. We have also

$$\operatorname{vol} S = e_{\text{in}}(S) + e_{\text{out}}(S).$$

A complete multipartite graph G is a graph whose vertices can be partitioned into pairwise disjoint subsets $V_1, \dots, V_k$ such that an edge exists if and only if its end vertices belong to different subsets. For $k = n$ we say that G is a complete graph, while for $k = 2$ and $|V_1| = 1$ we say that G is a star.

2. The community detection problem

The discovery and description of communities in a graph is a central problem in modern graph analysis. Intuitively, a community (or cluster) is a possibly connected group of nodes whose internal edges outnumber those with the rest of the network. However there is no formal definition of community. A survey of several recently proposed definitions can be found in [12], where the definition based on the modularity quality function is identified as a very relevant one. The modularity function was proposed by Newman and Girvan in [17] as a possible measure to quantify how much a subset $S \subset V$ is a "good cluster". They postulate that S is a cluster of nodes in G if the difference $Q(S)$ between the actual and the expected number of edges in the subgraph $G(S)$ is positive. The quantity $Q(S)$ is called *modularity* of S and is defined by the following equivalent formulas:

$$Q(S) = e_{\text{in}}(S) - \frac{(\operatorname{vol} S)^2}{\operatorname{vol} V} = \frac{\operatorname{vol} S \operatorname{vol} \bar{S}}{\operatorname{vol} V} - e_{\text{out}}(S). \tag{1}$$

Note the equalities $Q(S) = Q(\bar{S})$ and $Q(V) = 0$. The modularity of a vertex set is one of the most efficient indicators of its consistency as a community in G. For this reason, we adopt the following definition:

DEFINITION 2.1. A subgraph of G is a *module* if its vertex set S has positive modularity. If no ambiguity may occur, S is called a module itself.

The usefulness of the previous definition lies in the fact that, in practice, if $G(S)$ is a connected module whose size is significant then it can be recognized as a community.

Definition 2.1 leads naturally to an efficient measure of a partitioning of G into modules. Indeed, let $S_1, \dots, S_k$ be a partition of V into pairwise disjoint subsets. The *normalized modularity* of $S_1, \dots, S_k$ is defined as

$$q(S_1, \dots, S_k) = \frac{1}{\operatorname{vol} V} \sum_{i=1}^{k} Q(S_i). \tag{2}$$

The normalization factor $1/\operatorname{vol} V$ has been introduced in [15, 17] to settle the value of q in a range independent on G and k and for compatibility with previous works.

The problem of partitioning a graph into an arbitrary number of subgraphs whose overall modularity is maximized has received a considerable attention, not only in its applicative and computational aspects but also from the graph-theoretic point of view [6, 13]. The main contributions we propose in this work deal with the *cut version* of the community detection problem, that is the problem of finding a subset $S \subseteq V$ having maximal modularity. To this end, we define the *cut-modularity of the graph* G as the quantity

$$q_G^{Cut} = \max_{S \subseteq V} q(S, \bar{S}) = \frac{2}{\operatorname{vol} V} \max_{S \subseteq V} Q(S). \tag{3}$$

It is well known that the optimization of the modularity function (2) presents some drawbacks when employed for finding a partitioning of G into modules, since small clusters tend to be subsumed by larger ones. Among the many techniques and variants of the Newman–Girvan modularity that have been devised to tackle this issue, here we borrow from [1] two weighted versions of the modularity function that play a relevant role in the subsequent discussion:

- The *relative modularity* of $S \subseteq V$ is $Q_{\text{rel}}(S) = Q(S)/|S|$. This definition is naturally extended to the cut $\{S, \bar{S}\}$ as

$$q_{\text{rel}}(S, \bar{S}) = Q_{\text{rel}}(S) + Q_{\text{rel}}(\bar{S}) = Q(S)\frac{n}{|S||\bar{S}|}, \tag{4}$$

  which, in turn, leads to the definition of the *relative cut-modularity of* G

$$q_G^{RCut} = \max_{S \subseteq V} q_{\text{rel}}(S, \bar{S}).$$

- The *normalized modularity* of $S \subseteq V$ is defined as $Q_{\text{norm}}(S) = Q(S)/\operatorname{vol} S$ and that definition can be extended to the cut $\{S, \bar{S}\}$ as

$$q_{\text{norm}}(S, \bar{S}) = Q_{\text{norm}}(S) + Q_{\text{norm}}(\bar{S}) = Q(S)\frac{\operatorname{vol} V}{\operatorname{vol} S \operatorname{vol} \bar{S}}. \tag{5}$$

  As before we define the *normalized cut-modularity of the graph* G *as*

$$q_G^{NCut} = \max_{S \subseteq V} q_{\text{norm}}(S, \bar{S}).$$

Straightforward computations ensure

$$\frac{2\, q_G^{RCut}}{n\, d_{\max}} \leqslant q_G^{Cut} \leqslant \frac{q_G^{RCut}}{2}, \qquad \frac{2\, q_G^{NCut}}{\operatorname{vol} V} \leqslant q_G^{Cut} \leqslant \frac{q_G^{NCut}}{2}.$$

3. Modularity matrices and their properties

The probably best known methods for detecting a subset whose modularity well approximates the cut-modularity of G are based on the idea of spectral partitioning and are related with an important rank-one modification of the adjacency matrix, known as *the Newman–Girvan modularity matrix*. In analogy with graph Laplacians, in this section we define two different modularity matrices and describe a number of relevant structural properties.

3.1. The Newman–Girvan modularity matrix

Given a graph G and the associated adjacency matrix A, the *modularity matrix* of G has been introduced in [15] as

$$M = A - \frac{1}{\operatorname{vol} V} dd^T. \tag{6}$$

Note that we can express $Q(S)$ as

$$Q(S) = \mathbb{1}_S^T M \mathbb{1}_S. \tag{7}$$

The following proposition summarizes some basics properties of M:

PROPOSITION 3.1. *The matrix M satisfies the following properties:*

1. *M is symmetric and $M\mathbb{1} = 0$.*
2. *If $m_1 \geqslant \ldots \geqslant m_n$ are the eigenvalues of M and $\alpha_1 \geqslant \ldots \geqslant \alpha_n$ those of A then $\alpha_1 \geqslant m_1 \geqslant \alpha_2 \geqslant m_2 \geqslant \ldots \geqslant \alpha_n \geqslant m_n$.*
3. *0 is a simple eigenvalue of M if and only if A is nonsingular.*
4. *The largest eigenvalue of M is nonnegative, and is zero if and only if G is a complete multipartite graph.*

Proof. Point 1 is revealed by a direct computation. Point 2 is a direct consequence of the variational characterization of the eigenvalues of symmetric matrices, see e.g., [20]. To show point *3* we observe that the multiplicity of the zero eigenvalue of M is one plus the dimension of the kernel of A. Indeed consider the diagonal matrix $\Delta = \mathrm{Diag}(1/\sqrt{d_1}, \ldots, 1/\sqrt{d_n})$ and let $\delta = \Delta d$. Then $\Delta M \Delta \delta = 0$ and $\Delta A \Delta \delta = \delta$. Therefore the multiplicity of the zero eigenvalue of $\Delta M \Delta$ is the multiplicity of the zero eigenvalue of $\Delta A \Delta$ plus one. This proves point *3* as the multiplicity of 0 is invariant under matrix congruences. Point *4* is a rephrasing of [14, Thm. 1.1] and [2, Thm. 11]. $\square$

The modularity matrix M is at the basis of many spectral methods for community detection, and the eigenstructure of M can be used to describe clustering properties of graphs. A number of results relating algebraic properties of M to communities in G have appeared in recent literature [1, 2, 8, 9, 14]. As it often plays a special role in the algebraic analysis of the modular structure of G, the largest nonzero eigenvalue of M deserves a special symbol, borrowed from [8] and therein named *algebraic modularity*:

$$m_G = \max_{\substack{v \in \mathbb{R}^n \\ v^T \mathbb{1} = 0}} \frac{v^T M v}{v^T v}. \tag{8}$$

A major motivation behind spectral methods is the intuition that a close relation exists between m_G and the cut-modularity (3), and that the subsets having positive modularity are related with positive eigenvalues of M. The following theorem summarizes some important properties of M that have been proven in recent literature, supporting such intuition, see [2, 8, 14].

THEOREM 3.2. *Let $\lambda_i(M)$ denote the i-th largest eigenvalue of M. Then, the matrix M satisfies the following properties:*

1. *$m_G < \rho(A)$ and, if d is not an eigenvector of A then m_G is simple.*
2. *If G is not a complete graph or a complete multipartite graph then $m_G = \lambda_1(M) > 0$. If G is a star then $m_G = \lambda_2(M) < 0$. Otherwise (that is, if G is a complete graph or a complete multipartite graph which is not a star) $m_G = 0$.*
3. *$m_G \geqslant 2\langle d\rangle q_G^{Cut}$ where $\langle d\rangle = \mathrm{vol}V/n$ is the average degree of G.*
4. *Let $\{S_1,\dots,S_k\}$ be a partition that maximizes the quantity in (2), which has minimal cardinality, and which is made up entirely by modules. Then $k-1$ does not exceed the number of positive eigenvalues of M.*
5. *Let u be an eigenvector associated with m_G such that $d^Tu \geqslant 0$. If m_G is simple and it is not an eigenvalue of A then the subgraph induced by the subset $S_+ = \{i \mid u_i \geqslant 0\}$ is connected.*

For any $S \subseteq V$ let $v_S = \mathbb{1}_S - \frac{|S|}{n}\mathbb{1}$. The following identities are readily obtained:

$$v_S^T\mathbb{1} = 0, \qquad v_S^Tv_S = \frac{|S||\bar S|}{n}, \qquad v_S^TMv_S = Q(S), \qquad q_{\mathrm{rel}}(S,\bar S) = \frac{v_S^TMv_S}{v_S^Tv_S}.$$

Hence, the combinatorial problem of finding the cut $\{S,\bar S\}$ with largest relative modularity has a natural continuous relaxation in the maximization of the Rayleigh quotient v^TMv/v^Tv over the subspace orthogonal to $\mathbb{1}$, that is, the algebraic modularity defined in (8). We have the immediate consequence

$$q_G^{RCut} \leqslant m_G.$$

3.2. The normalized modularity matrix

In analogy with the normalized Laplacian matrix of a graph, we define the *normalized modularity matrix* of G as

$$\mathscr{M} = D^{-1/2}MD^{-1/2} = \mathscr{A} - \frac{1}{\mathrm{vol}V}\delta\delta^T$$

where $\delta = (\sqrt{d_1},\dots,\sqrt{d_n})^T$ and M is as in (6). The matrix $\mathscr{M}$ appeared recently in the community detection literature, and in various other network related questions as the analysis of quasi-randomness properties of graphs with given degree sequences, see [1, 4, 9] and [3, Chap. 5]. Several basics properties of $\mathscr{M}$ can be immediately observed; we collect some of them hereafter.

PROPOSITION 3.3. *The matrix $\mathscr{M}$ satisfies the following properties:*

1. *$\mathscr{M}$ has a zero eigenvalue with corresponding eigenvector δ.*

2. *$\mathscr{M}v = \mathscr{A}v$ for all vectors v orthogonal to δ.*

3. *The eigenvalues of $\mathscr{M}$ belong to the interval $[-1,1]$. Moreover, 0 is a simple eigenvalue of $\mathscr{M}$ if and only if $\mathscr{A}$ is nonsingular.*

4. *If G is connected then 1 is not an eigenvalue of $\mathscr{M}$. Furthermore, if G is not bipartite then -1 is not an eigenvalue of $\mathscr{M}$.*

Proof. Straightforward computations show that $\mathscr{A}\delta = \delta$ and $\mathscr{M}\delta = 0$. Since $\mathscr{A} \geqslant O$ and $\delta \geqslant 0$, Perron–Frobenius theory leads us to deduce that $\rho(\mathscr{A}) = 1$ is an eigenvalue of $\mathscr{A}$. Therefore, if $\mathscr{A} = \sum_{i=1}^n \lambda_i q_i q_i^T$ is a spectral decomposition of $\mathscr{A}$ with the eigenvalues in nonincreasing order, $\lambda_1 \geqslant \ldots \geqslant \lambda_n$, then we can assume $\lambda_1 = 1$, $|\lambda_i| \leqslant 1$ for $i > 1$, and q_1 parallel to δ. In particular, $\delta\delta^T/\mathrm{vol}\,V$ is the orthogonal projector on the eigenspace spanned by q_1, since $\delta^T\delta = \mathrm{vol}\,V$. Consequently, $\mathscr{M} = \sum_{i=2}^n \lambda_i q_i q_i^T$ is a spectral decomposition of $\mathscr{M}$ and we easily deduce points *2* and *3*. Incidentally, this proves that $\mathscr{M}$ and $\mathscr{A}$ are simultaneously diagonalizable. If G is connected then $\mathscr{A}$ is irreducible and λ_1 is simple, that is $1 > \lambda_2$. Furthermore, if G is not bipartite then $\mathscr{A}$ is also primitive and $|\lambda_i| < 1$ for $i > 1$, and the proof is complete. $\square$

The normalized modularity (5) of a cut $\{S, \bar{S}\}$ can be naturally defined in terms of $\mathscr{M}$. In fact, given any $S \subseteq V$, consider the vector

$$v_S = D^{1/2}(\mathbb{1}_S - c\mathbb{1}), \qquad c = \mathrm{vol}\,S/\mathrm{vol}\,V. \tag{9}$$

Simple computations prove that

$$\delta^T v_S = 0, \qquad v_S^T v_S = \frac{\mathrm{vol}\,S\,\mathrm{vol}\,\bar{S}}{\mathrm{vol}\,V}.$$

Moreover,

$$\frac{v_S^T \mathscr{M} v_S}{v_S^T v_S} = \frac{(\mathbb{1}_S - c\mathbb{1})^T M (\mathbb{1}_S - c\mathbb{1})}{v_S^T v_S} = \frac{\mathbb{1}_S^T M \mathbb{1}_S}{\mathrm{vol}\,S\,\mathrm{vol}\,\bar{S}}\mathrm{vol}\,V = q_{\mathrm{norm}}(S, \bar{S}).$$

It follows that the problem of computing the normalized cut-modularity of G can be stated in terms of $\mathscr{M}$. Indeed, if $\mathscr{V}_n$ is the set of n-vectors having the form (9) for some $S \subset V$, then

$$q_G^{NCut} = \max_{v \in \mathscr{V}_n} \frac{v^T \mathscr{M} v}{v^T v} \tag{10}$$

and of course, if $\hat{v}$ is the vector realizing the maximum in (10), then the set $\hat{S} = \{i \mid \hat{v}_i > 0\}$ defines the optimal cut. As for the unnormalized case, we define the *normalized algebraic modularity*:

$$\mu_G = \max_{\substack{v \in \mathbb{R}^n \\ v^T\delta = 0}} \frac{v^T \mathscr{M} v}{v^T v}. \tag{11}$$

Note that (11) is a relaxed version of (10). In particular,

$$q_G^{NCut} \leqslant \mu_G. \tag{12}$$

Since $\mathscr{M}$ is real symmetric, μ_G coincides with the largest eigenvalue of $\mathscr{M}$ after deflation of the invariant subspace spanned by δ. Therefore, if $-1 \leqslant \mu_n \leqslant \cdots \leqslant \mu_1 \leqslant 1$ are the eigenvalues of $\mathscr{M}$, then $\mu_1 = \max\{0, \mu_G\}$. Furthermore, since M and $\mathscr{M}$ are congruent matrices, point *2* of Theorem 3.2 leads us to the following result:

COROLLARY 3.4. *If G is not a star then $\mu_G = \mu_1$, the largest eigenvalue of $\mathscr{M}$. Moreover, $\mu_G > 0$ if and only if G is not a complete graph or a complete multipartite graph.*

4. Cheeger-type inequalities

As we already discussed above, both heuristics and intuition suggest that μ_G quantifies the cut-modularity of the graph, and can be used to approximate q_G^{NCut}. While the upper bound $q_G^{NCut} \leqslant \mu_G$ has been shown in (12) by simple arguments, a converse relation, bounding q_G^{NCut} from below in terms of μ_G, is not that easy. In fact, it is possible that $\mu_G > 0$ while $q_G^{NCut} < 0$, as shown experimentally in [2]. Theorems 4.1 and 4.3 contribute to this question stating lower (and upper) bounds of q_G^{NCut} in terms of spectral properties of $\mathscr{M}$.

The conductance (or sparsity, or Cheeger constant) h_G is one of the best known topological invariants of a graph G, defined as follows: For $S \subset V$ let

$$h(S) = \frac{e_{\text{out}}(S)}{\min\{\operatorname{vol} S, \operatorname{vol} \bar{S}\}}$$

and $h_G = \min_{S \subset V} h(S)$. Such quantity plays a fundamental role in graph partitioning problems [16, Chap. 11], in isoperimetric problems [3, Chap. 2], mixing properties of random walks, combinatorics, and in various other areas of mathematics and computer science. A renowned result in graph theory, known as *Cheeger inequality*, relates the conductance of G and the smallest positive eigenvalue of the normalized Laplacian matrix $\mathscr{L} = I - \mathscr{A}$.

If $0 = \lambda_1 < \lambda_2 \leqslant \ldots \leqslant \lambda_n \leqslant 2$ are the eigenvalues of $\mathscr{L}$, the Cheeger inequality states that

$$\tfrac{1}{2}\lambda_2 \leqslant h_G \leqslant \sqrt{2\lambda_2}.$$

Chung [3] improved the upper bound to $h_G \leqslant \sqrt{\lambda_2(2-\lambda_2)}$. Let v be an eigenvector of $\mathscr{L}$ corresponding to λ_2 and consider the equality $\mathscr{L} = I - \mathscr{A} = I - \mathscr{M} + \delta\delta^T/\delta^T\delta$. Since $\mathscr{L}\delta = 0$, we have $\delta^T v = 0$. By Courant's minimax principle and (11),

$$\lambda_2 = \min_{v:\delta^T v = 0} \frac{v^T \mathscr{L} v}{v^T v} = 1 - \max_{v:\delta^T v = 0} \frac{v^T \mathscr{M} v}{v^T v} = 1 - \mu_G.$$

In particular, from Corollary 3.4 we obtain that, if G is not a star then $1 - \lambda_2$ is the largest eigenvalue of $\mathscr{M}$. A direct application of the Cheeger inequality yields the following estimates for q_G^{NCut}.

THEOREM 4.1. *Let μ_1 be the largest eigenvalue of $\mathscr{M}$. If G is not a star then*

$$1-2\sqrt{1-\mu_1^2} \leqslant q_G^{NCut} \leqslant \mu_1.$$

Proof. Recalling (1) and (5), we have

$$\begin{aligned} q_{\text{norm}}(S,\bar{S}) &= \frac{\operatorname{vol} V}{\operatorname{vol} S \operatorname{vol} \bar{S}} Q(S) \\ &= 1 - \frac{\operatorname{vol} V}{\operatorname{vol} S \operatorname{vol} \bar{S}} e_{\text{out}}(S) \geqslant 1 - 2h(S), \end{aligned}$$

since $\operatorname{vol} V/\operatorname{vol} S \operatorname{vol} \bar{S} \leqslant 2/\min\{\operatorname{vol} S, \operatorname{vol} \bar{S}\}$. By maximizing over S we eventually get

$$q_G^{NCut} = \max_{S\subset V} q_{\text{norm}}(S,\bar{S}) \geqslant 1-2h_G \geqslant 1-2\sqrt{(1-\mu_G)(1+\mu_G)}.$$

By hypothesis, $\mu_G = \mu_1$. The upper bound comes from (12). □

Extensive research on Cheeger-type results by many authors suggests that no substantial improvements on the lower bound in Theorem 4.1 can be obtained without additional information on G, although explicit examples of graph sequences proving optimality of that bound are not known. However, the forthcoming result shows that $1-\mu_1$ can be a much better estimate to $1-q_G^{NCut}$ than expected when the entries of an eigenvector of μ_1 cluster around two values. We will make use of the following lemma, whose simple proof is omitted for brevity:

LEMMA 4.2. *If $\sum_{i=1}^n \alpha_i = 0$ then $\sum_{i:\alpha_i>0} \alpha_i = \frac{1}{2}\sum_{i=1}^n |\alpha_i|$.*

THEOREM 4.3. *Let μ_1 be the largest eigenvalue of $\mathscr{M}$. Suppose that μ_1 has an eigenvector x without zero entries. Then there exists a constant $C > 0$, not depending on μ_1, such that*

$$1 - C(1-\mu_1) \leqslant q_G^{NCut}.$$

Proof. Let v be an eigenvector of $\mathscr{M}$ corresponding to μ_1 and let $z = D^{-1/2}v$. Note that v is orthogonal to the vector $\delta = (\sqrt{d_1},\dots,\sqrt{d_n})^T$, since the latter is an eigenvector of $\mathscr{M}$ associated to 0. Consequently, z is orthogonal to the degree vector: $d^Tz = \delta^T D^{1/2} z = \delta^T v = 0$. Hence,

$$\mu_1 = \frac{v^T \mathscr{M} v}{v^T v} = \frac{v^T \mathscr{A} v}{v^T v} = \frac{z^T A v}{z^T D z} = 1 - \frac{z^T L z}{z^T D z},$$

where $L = D - A$ is the Laplacian matrix of G. We have

$$z^T L z = \sum_{ij\in E} (z_i - z_j)^2,$$

where the sum runs over the edges of the graph, each edge being counted only once. On the other hand,

$$z^T D z = \sum_{i=1}^n d_i z_i^2 .$$

For notational simplicity, let $s = \mathrm{vol}\, S$, $\bar{s} = \mathrm{vol}\, \bar{S}$, and $\nu = s + \bar{s} = \mathrm{vol}\, V$. Consider the nodal domain $S = \{i : v_i \geqslant 0\}$ and let x be the vector $x = p\mathbb{1}_S + q\mathbb{1}_{\bar{S}}$ which minimizes the weighted distance

$$\|D^{1/2}(x-z)\|_2^2 = \sum_{i=1}^n d_i(x_i - z_i)^2 = \sum_{i\in S} d_i(p-z_i)^2 + \sum_{i\in\bar{S}} d_i(q-z_i)^2 .$$

Simple computations show that the minimum is attained when

$$p = \Big(\sum_{i\in S} d_i z_i\Big)/s, \qquad q = \Big(\sum_{i\in\bar{S}} d_i z_i\Big)/\bar{s}.$$

Observe that p and q are weighted averages of the values z_i for $i \in S$ and $i \in \bar{S}$, respectively. With the notation $c = \sum_{i\in S} d_i z_i$, from the orthogonality condition $d^T z = 0$ and Lemma 4.2 we deduce that $p = c/s$ and $q = -c/\bar{s}$. For later reference, we remark the identities

$$p - q = \frac{c\nu}{s\bar{s}}, \qquad p^2 s + q^2 \bar{s} = \nu \frac{(c\nu)^2}{(s\bar{s})^2}. \tag{13}$$

Note that, apart of a constant, the vector $D^{1/2}x$ coincides with the vector in (9). It is not hard to recognize that, if G is disconnected then the vector $D^{1/2}x$ is an eigenvector of $\mathscr{M}$ associated to the eigenvalue 1. Our subsequent arguments are based on the intuition that, if z is a small perturbation of x then S is weakly linked to $\bar{S}$. Let $r \geqslant 1$ be a number such that

$$r^{-1} \leqslant z_i/x_i \leqslant r, \qquad i = 1,\dots,n.$$

In fact, if $z_i > 0$ then $x_i = p > 0$, whereas $z_i < 0$ implies $x_i = q < 0$. Hence, if $ij \in E$ is an edge joining a node in S with a node in $\bar{S}$ we have $|z_i - z_j| \geqslant (p-q)/r$. Consequently,

$$z^T L z = \sum_{ij\in E} (z_i - z_j)^2 \geqslant r^{-2}(p-q)^2 e_{\mathrm{out}}(S),$$

by neglecting all contributions from edges lying entirely inside S or $\bar{S}$. Moreover,

$$z^T D z = \sum_{i=1}^n d_i z_i^2 \leqslant r^2\Big(\sum_{i\in S} p^2 d_i + \sum_{i\in\bar{S}} p^2 d_i\Big) = r^2(p^2 s + q^2\bar{s}).$$

Consider the equality $e_{\mathrm{out}}(S) = (1 - q_{\mathrm{norm}}(S,\bar{S}))s\bar{s}/\nu$. Using (13) and simplifying we get

$$1 - \mu = \frac{z^T L z}{z^T D z} \geqslant \frac{1}{r^4 \nu} e_{\mathrm{out}}(S) = \frac{s\bar{s}}{r^4\nu^2}(1 - q_{\mathrm{norm}}(S,\bar{S})) \geqslant \frac{1}{4r^4}(1 - q_G^{NCut}),$$

owing to $s\bar{s}/\nu^2 \geqslant \frac{1}{4}$. $\square$

5. Modules from nodal domains

Theorems 4.1 and 4.3 prove that if μ_G is sufficiently close to 1 then the cut-modularity of G is positive and thus there exists a bipartition of V into $\{S,\bar{S}\}$ such that both $G(S)$ and $G(\bar{S})$ are modules. Of course such bipartition is not unique in the general case. The forthcoming theorems strengthen this claim by showing that, if a positive eigenvalue μ of $\mathscr{M}$ is large enough, then we can explicitly exhibit a cut $\{S,\bar{S}\}$ with positive modularity, by defining it in terms of a nodal domain induced by an eigenvector corresponding to μ.

Given a nonzero vector $v\in\mathbb{R}^n$ the subgraph $G(S)$ induced by the set $S=\{i : v_i\geqslant 0\}$ is a *nodal domain* of v [5, 7]. This fundamental definition admits obvious variations (for example, inequality can be strict, or reversed) and, since the seminal papers by Fiedler [10, 11], it has become a major tool for spectral methods in community detection and graph partitioning [15, 18, 19]. If v is an eigenvector corresponding to μ_G, it has been shown in [9] that $S=\{i : v_i\geqslant 0\}$ induces a connected subgraph $G(S)$. The following Theorems 5.1 and 5.2 provide additional information on $G(S)$ as they show that, if μ_G is large enough, then the subgraph $G(S)$ is a module.

THEOREM 5.1. *Let v be a normalized eigenvector of $\mathscr{M}$ corresponding to a positive eigenvalue μ, that is, $\mathscr{M}v=\mu v$ with $\|v\|_2=1$. Let $S=\{i \mid v_i\geqslant 0\}$. If*

$$\mu > \frac{(\operatorname{vol}S)^2+(\operatorname{vol}\bar{S})^2}{\operatorname{vol}V}\max_{i\in V}\frac{v_i^2}{d_i}$$

then $Q(S)>0$.

Proof. Recalling Proposition 3.3, we have that v is orthogonal to δ, which implies in turn $\mathscr{M}v=\mathscr{A}v$ and $\mu=v^T\mathscr{M}v=v^T\mathscr{A}v$. Define the set $\mathscr{I}_+=(S\times S)\cup(\bar{S}\times\bar{S})$. Note that $v_iv_j\geqslant 0$ whenever $(i,j)\in\mathscr{I}_+$. Using entrywise nonnegativity of $\mathscr{A}$ we obtain

$$\mu=v^T\mathscr{A}v\leqslant\sum_{(i,j)\in\mathscr{I}_+}v_iv_j\mathscr{A}_{ij}\leqslant\left(\max_{i\in V}\frac{|v_i|}{\delta_i}\right)^2\sum_{(i,j)\in\mathscr{I}_+}\delta_i\delta_j\mathscr{A}_{ij}.$$

Since $\delta_i\delta_j\mathscr{A}_{ij}=A_{ij}$, the rightmost summations yield

$$\sum_{(i,j)\in\mathscr{I}_+}A_{ij}=\mathbb{1}_S^TA\mathbb{1}_S+\mathbb{1}_{\bar{S}}^TA\mathbb{1}_{\bar{S}}=e_{\text{in}}(S)+e_{\text{in}}(\bar{S}).$$

Let us set $C^2=(\max_{i\in V}|v_i|/\delta_i)^2$. Owing to the equalities $Q(S)=e_{\text{in}}(S)-(\operatorname{vol}S)^2/\operatorname{vol}V$ and $Q(S)=Q(\bar{S})$ we have

$$\mu\leqslant C^2\big(e_{\text{in}}(S)+e_{\text{in}}(\bar{S})\big)=C^2\Big(2Q(S)+\frac{(\operatorname{vol}S)^2+(\operatorname{vol}\bar{S})^2}{\operatorname{vol}V}\Big).$$

By rearranging terms,

$$2C^2Q(S)\geqslant\mu-C^2\frac{(\operatorname{vol}S)^2+(\operatorname{vol}\bar{S})^2}{\operatorname{vol}V},$$

and the claim follows. $\square$

With respect to the quantity $\max_i v_i^2/d_i$ appearing in the preceding theorem, consider that if G is k-regular (that is, $d_i = k$ for every $i \in V$) then $v_i = n^{-\frac{1}{2}}$ and $\mathrm{vol}\, V = kn$. After simple passages the lower bound for μ becomes $(|S|^2 + |\bar{S}|^2)/n^2$, a number which is strictly smaller than 1.

THEOREM 5.2. *Let v be any real eigenvector of $\mathscr{M}$ corresponding to a positive eigenvalue μ, that is, $\mathscr{M} v = \mu v$. Let $S = \{i \mid v_i \geqslant 0\}$ and let $\cos\theta$ be the cosine of the acute angle between the vectors $|v| = (|v_1|, \dots, |v_n|)^T$ and $\delta = (\sqrt{d_1}, \dots, \sqrt{d_n})^T$. If*

$$\mu + 1 > 4 \frac{\mathrm{vol}\, S \,\mathrm{vol}\, \bar{S}}{(\mathrm{vol}\, V)^2} \frac{1}{\cos^2\theta}$$

then $Q(S) > 0$.

Proof. Let $s = D^{1/2} \mathbb{1}_S$, that is

$$s_i = \begin{cases} \delta_i & v_i \geqslant 0, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $\|s\|_2^2 = \sum_{i \in S} d_i = \mathrm{vol}\, S$ and $\delta^T s = \mathrm{vol}\, S$ too. Since $\delta^T v = 0$ there exist scalars α, β, γ such that we have the orthogonal decomposition

$$s = \alpha \frac{1}{\|\delta\|_2} \delta + \beta \frac{1}{\|v\|_2} v + \gamma w \tag{14}$$

for some normalized vector $w \in \mathbb{R}^n$ orthogonal to both δ and v. The coefficients in (14) own the following explicit formulas:

$$\alpha = \frac{1}{\|\delta\|_2} \delta^T s = \frac{\mathrm{vol}\, S}{\sqrt{\mathrm{vol}\, V}}, \qquad \beta = \frac{v^T s}{\|v\|_2},$$

and moreover,

$$\begin{aligned} \gamma^2 &= \|s\|_2^2 - \alpha^2 - \beta^2 = \mathrm{vol}\, S - \frac{(\mathrm{vol}\, S)^2}{\mathrm{vol}\, V} - \beta^2 \\ &= \frac{\mathrm{vol}\, S \,\mathrm{vol}\, \bar{S}}{\mathrm{vol}\, V} - \beta^2. \end{aligned}$$

Owing to the fact that the spectrum of $\mathscr{M}$ is included in $[-1, 1]$ and the assumption $\|w\|_2 = 1$ we have $w^T \mathscr{M} w \geqslant -1$. Hence, from (14) we obtain

$$\begin{aligned} Q(S) &= \mathbb{1}_S^T M \mathbb{1}_S = s^T \mathscr{M} s \\ &\geqslant \alpha^2 \cdot 0 + \beta^2 \mu - \gamma^2 = \beta^2(\mu + 1) - \frac{\mathrm{vol}\, S \,\mathrm{vol}\, \bar{S}}{\mathrm{vol}\, V}. \end{aligned}$$

Thus, if

$$\mu + 1 > \frac{\operatorname{vol} S \operatorname{vol} \overline{S}}{\beta^2 \operatorname{vol} V}$$

then $Q(S) > 0$. Moreover, from $\delta^T v = 0$ and Lemma 4.2 we obtain

$$\cos\theta = \frac{\sum_{i\in V} \delta_i |v_i|}{\|v\|_2 \|\delta\|_2} = \frac{2\sum_{i\in S} \delta_i v_i}{\|v\|_2 \sqrt{\operatorname{vol} V}} = 2\frac{v^T s}{\|v\|_2 \sqrt{\operatorname{vol} V}},$$

whence $\beta = \frac{1}{2}(\cos\theta)\sqrt{\operatorname{vol} V}$ and the proof is complete. □

From the straightforward bound

$$\operatorname{vol} S \operatorname{vol} \overline{S} / (\operatorname{vol} V)^2 \leqslant \tfrac{1}{4}$$

and the equality $\cos^{-2}\theta - 1 = \tan^2\theta$, we derive the following condition.

COROLLARY 5.3. *In the same notations of Theorem 5.2, if* $\mu > \tan^2\theta$ *then* $Q(S) > 0$.

6. Concluding remarks

Community detection is a major task in modern complex network analysis and the matrix approach to such problem is quite popular and powerful. In this work we formulate the modularity of a cut in terms of a quadratic form associated with the normalized modularity matrix, and we provide theoretical supports to the common understanding that highly positive eigenvalues of the normalized modularity matrix imply the presence of communities in G. In particular we show that, if that matrix has an eigenvalue close to 1 then the nodal domains corresponding to that eigenvalue have positive modularity and, moreover, can produce good estimates of the optimal cut-modularity.

Acknowledgement. The first author acknowledges the support received by Istituto Nazionale di Alta Matematica (INdAM-GNCS, Italy) for his research. The work of the second author is partially supported by the ERC grant NOLEPRO.

REFERENCES

[1] M. BOLLA, *Penalized versions of the Newman–Girvan modularity and their relation to normalized cuts and k-means clustering*, Phys. Rev. E, **84** (2011), 016108.

[2] M. BOLLA, B. BULLINS, S. CHATURAPRUEK, S. CHEN, AND K. FRIEDL, *Spectral properties of modularity matrices*, Linear Algebra Appl., **473** (2015), 359–376.

[3] F. CHUNG, *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics, **92**, AMS, Providence, 1997.

[4] F. CHUNG AND R. GRAHAM, *Quasi-random graphs with given degree sequences*, Random Structures Algorithms, **32** (2008), 1–19.

[5] E. B. DAVIES, G. M. L. GLADWELL, J. LEYDOLD, AND P. F. STADLER, *Discrete nodal domain theorems*, Linear Algebra Appl., **336** (2001), 51–60.

[6] F. DE MONTGOLFIER, M. SOTO, AND L. VIENNOT, *Asymptotic modularity of some graph classes*, In: Algorithms and computation, vol. 7074 of Lecture Notes in Comput. Sci., pages 435–444. Springer, Heidelberg, 2011.

[7] A. M. DUVAL AND V. REINER, *Perron–Frobenius type results and discrete versions of nodal domain theorems*, Linear Algebra Appl., **294** (1999), 259–268.

[8] D. FASINO AND F. TUDISCO, *An algebraic analysis of the graph modularity*, SIAM J. Matrix Anal. Appl., **35** (2014), 997–1018.

[9] D. FASINO AND F. TUDISCO, *Generalized modularity matrices*, Linear Algebra Appl., **502** (2016), 327–345.

[10] M. FIEDLER, *Algebraic connectivity of graphs*, Czechoslovak Mathematical Journal, **23** (1973), 298–305.

[11] M. FIEDLER, *A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory*, Czechoslovak Mathematical Journal, **25** (1975), 619–633.

[12] S. FORTUNATO, *Community detection in graphs*, Physics Reports, **486** (2010), 75–174.

[13] A. KEHAGIAS AND L. PITSOULIS, *Bad communities with high modularity*, Eur. Phys. J. B, **86** (2013), Art. 330.

[14] S. MAJSTOROVIC AND D. STEVANOVIC, *A note on graphs whose largest eigenvalues of the modularity matrix equals zero*, Electron. J. Linear Algebra, **27** (2014), 611–618.

[15] M. E. J. NEWMAN, *Finding community structure in networks using the eigenvectors of matrices*, Phys. Rev. E, **74** (2006), 036104.

[16] M. E. J. NEWMAN, *Networks: An Introduction*, OUP Oxford, Oxford, 2010.

[17] M. E. J. NEWMAN AND M. GIRVAN, *Finding and evaluating community structure in networks*, Phys. Rev. E, **69** (2004), 026113.

[18] D. L. POWERS, *Graph partitioning by eigenvectors*, Linear Algebra Appl., **101** (1988), 121–133.

[19] S. E. SCHAEFFER, *Graph clustering*, Computer Science Review, **1** (2007), 27–64.

[20] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, 1965.

(Received October 28, 2015)

Dario Fasino
Department of Mathematics, Computer Science and Physics
University of Udine
Udine, Italy
e-mail: dario.fasino@uniud.it

Francesco Tudisco
Department of Mathematics
University of Padua
Padua, Italy
e-mail: francesco.tudisco@math.unipd.it