Spec. Matrices 2018; 6:110–121 DE GRUYTER

Research Article Open Access

Dario Fasino and Francesco Tudisco*

The expected adjacency and modularity matrices in the degree corrected stochastic block model

https://doi.org/10.1515/spma-2018-0010
Received October 20, 2017; accepted February 16, 2018

Abstract: We provide explicit expressions for the eigenvalues and eigenvectors of matrices that can be written as the Hadamard product of a block partitioned matrix with constant blocks and a rank one matrix. Such matrices arise as the expected adjacency or modularity matrices in certain random graph models that are widely used as benchmarks for community detection algorithms.

Keywords: Adjacency matrix, modularity matrix, stochastic block model, inflation product.

MSC: 15A18, 15B99

1 Introduction

Graph clustering is one of the most relevant tasks in the structural analysis of complex networks. In a broad sense, it consists of detecting and investigating various types of *meso-scale* structures, typically identified with groups of nodes showing distinctive forms of relationships in their interior and with the rest of the network. For instance, group of nodes being rich of internal connections but being loosely connected with outside nodes are indication of communities. Communities have received a considerable attention in recent years because they emerge naturally in graphs representing "assortative" relations, namely, those relations that are established preferably among entities that are in some sense similar, as for instance in social networks with friendship relations or in data sets turned into graphs via spatial distances or other similarity measures [26].

However, other situations are possible and are receiving growing attention in the network science literature. For example, when the graph encodes "disassortative" relations, such as food webs including predator and prey species or social networks with ties representing hostilities, an important task is the identification of anti-communities, being groups of nodes with many connections between different groups but having relatively loose internal connections [6]. Moreover, communities and anti-communities may coexist, for instance in communication, recommendation, and collaboration networks [10, 11, 19].

All these network structures are often well described by means of block models that describe idealized situations [9]. Figure 1 shows a pictorial representation of some meso-scale structures that can be effectively described by block models. These ideal block structures are simple to analyze, give a reasonable approximation of structured real-world graphs and are easy to generate "artificially". The Stochastic Block Model, together with its degree corrected variant [5, 8, 23], is one of the most successful and widespread generative models for graphs showing meso-scale block structures. This random network model is a generalization of various random graph models, including the best known Erdös–Rényi, Chung–Lu, and the planted partition

Dario Fasino: Dipartimento di Scienze Matematiche, Informatiche e Fisiche, Università di Udine, Udine, Italy. E-mail: dario.fasino@uniud.it
***Corresponding Author: Francesco Tudisco:** Department of Mathematics and Statistics, University of Strathclyde, Glasgow, UK. E-mail: f.tudisco@strath.ac.uk

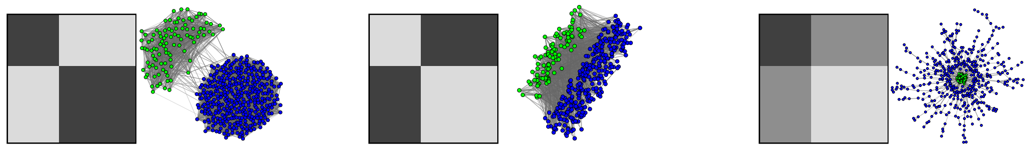

Figure 1: Examples of block models for meso-scale network structures. The squares on the left represent densities of non-zero entries in idealized adjacency matrices. The structures here shown are (from left to right): *community*, *anti-community*, *core-periphery*.

models, where a ground truth clustering assignment is defined a-priori on the node set. The latter property is a particularly relevant aspect of the model as it allows for quantitative evaluations of graph clustering algorithms, where a number of important graph matrices play a crucial role. In fact, many clustering algorithms are grounded on the spectral analysis of graph matrices, as for instance the adjacency [1, 24], the Laplacian [26, 28], the signless Laplacian [21], the Bhete Hessian [25], etc. The analysis of (spectral) matrix-based clustering methods for networks following stochastic block models often helps shedding light on the performances of different methods and, by looking at the graphs in expectation, allows to prove a number of quality guarantees holding with high probability. It is worth recalling that graph clustering is in general a provably hard problem, as any precise mathematical formulation can only be given in terms of combinatorial optimization [9, 26]; on the other hand, it is well known that spectral techniques often outperform more traditional clustering algorithms [28].

Here we carry out a thorough spectral analysis of the expected adjacency matrix belonging to the degree corrected Stochastic Block Model (dc-SBM). Our analysis is motivated by the occurrence of the dc-SBM and the spectral properties of graph matrices in expectation under such model for the theoretical analysis of spectral algorithms for clustering and community detection, see e.g., [5, 8, 21, 23, 24]. For definiteness, we will make reference to the following prototypical argument throughout this paper: the adjacency matrix A of a random graph drawn from a dc-SBM can be decomposed into the sum of a constant matrix $\overline{A}$ and a matrix with random entries W, called Wigner noise [1]. The matrix $\overline{A}$ is the expected adjacency matrix of the model, and its spectral analysis makes available enough information to recover all the parameters defining the model. If the spectra of $\overline{A}$ and W are well separated, standard tools in perturbation analysis allow to relate the eigendecomposition of A to that of $\overline{A}$. Consequently, algorithms based on the computation of a few eigenpairs of A can provide reliable indications on the meso-scale structure of a graph belonging to a stochastic block model.

In this work we focus also on the modularity matrix, a further widely explored graph matrix, defined as a rank-one modification of the adjacency matrix [22], whose usage is very successful in various clustering problems. Important algebraic properties of this matrix have been proved in recent years [3, 12, 14, 20], and in [13] we provided theoretical evidence to the fact that clusters located by leading eigenvectors of modularity matrices can be recognized as communities inside the graph.

By means of the matrix inflation operator [15], we reveal that expected adjacency and modularity matrices of stochastic block models enjoy a particular block structure with rank structured blocks. That characterization allows us to obtain explicit formulas for the *structural* eigenvalues and eigenvectors of these matrices, that is, the eigenpairs providing information on the model parameters. Furthermore, we describe precise asymptotic properties for the structural eigenvalues of sequences of expected adjacency matrices with growing dimensions, under very general hypotheses on the model structure.

Notation

In the following we use standard graph and matrix theoretic notations and concepts [10, 17]. In particular, we denote by **1** the vector of all ones, the dimension of such vector being given by the context and, for a vector $\boldsymbol{x} = (x_1, \ldots, x_m)^T \in \mathbb{R}^m$, we use interchangeably the symbols $\mathrm{Diag}(\boldsymbol{x})$ and $\mathrm{Diag}(x_1, \ldots, x_m)$ to denote the $m \times m$ diagonal matrix with diagonal entries being the entries of $\boldsymbol{x}$.

The adjacency matrix of an undirected graph on the node set $\{1, \dots, n\}$ is the symmetric $n \times n$ matrix $A = (a_{ij})$ such that $a_{ij} = 1$ if ij is an edge of the graph and $a_{ij} = 0$ otherwise. We allow loops, that is, edges joining a node with itself. The associated Newman–Girvan modularity matrix is the matrix $M = A - \boldsymbol{d}\boldsymbol{d}^T/\nu$ where $\boldsymbol{d} = A\mathbf{1}$ is the degree vector and $\nu = \boldsymbol{d}^T\mathbf{1}$ is the volume of the graph, that is, the sum of all node degrees [22]. That matrix, along with some variants and generalizations, plays a leading role in the design of spectral methods for community detection. Note that the vector $\mathbf{1}$ belongs to the kernel of M:

$$M\mathbf{1} = A\mathbf{1} - \boldsymbol{d}(\boldsymbol{d}^T\mathbf{1})/\nu = \boldsymbol{d} - \boldsymbol{d} = 0.$$

2 The stochastic block model

The Stochastic Block Model (SBM) is one of the most widespread generative models for random graphs in the graph clustering literature. An SBM with n nodes and k blocks is a random graph model parametrized by the *membership matrix* $\Theta \in \{0, 1\}^{n\times k}$ and the symmetric *connectivity matrix* $B = (b_{ij}) \in \mathbb{R}^{k\times k}$. Every row of the matrix Θ contains exactly one nonzero entry, whose position indicates which block that node belongs to; and the nonzero entries in the i-th column indicate nodes belonging to the i-th block. For $i = 1, \dots, n$, we denote by $\beta(i)$ the block index of node i. Hence, $\Theta_{ij} = 1$ iff $j = \beta(i)$.

The entry b_{ij} is the edge probability between any node in block i and any node in block j. That is, for any two nodes p, q, the probability that they are joined by an edge is $b_{\beta(p),\beta(q)}$. Edges are generated independently from one another.

We denote by n_i the number of elements in the i-th block, $n_1 + \cdots + n_k = n$. For visual convenience, we assume that each block consists of consecutive integers. Hence, the expected adjacency matrix within the SBM with parameters (Θ, B) is naturally partitioned into the block form

$$\overline{A} = \begin{pmatrix} \overline{A}_{11} & \cdots & \overline{A}_{1k} \\ \vdots & & \vdots \\ \overline{A}_{k1} & \cdots & \overline{A}_{kk} \end{pmatrix}, \qquad \overline{A}_{ij} = b_{ij}\mathbf{1}\mathbf{1}^T \in \mathbb{R}^{n_i\times n_j}, \tag{1}$$

and can be expressed in factored form as $\overline{A} = \Theta B \Theta^T$.

2.1 The degree corrected SBM

The SBM has one major drawback that makes it unsuitable for modeling networks found in the real world, that is, it presumes that all nodes within the same block have the same expected degree. The *degree corrected stochastic block model* is a variant of SBM designed to allow a greater degree heterogeneity [5, 8, 23]. A degree corrected SBM (dc-SBM) is defined by a triple $(\Theta, B, \boldsymbol{\delta})$ where Θ and B are as in the SBM and $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)^T$ is a vector whose entries are positive real numbers. The expected adjacency matrix in this model is defined as

$$\overline{A} = \mathrm{Diag}(\boldsymbol{\delta})\Theta B \Theta^T \mathrm{Diag}(\boldsymbol{\delta})$$

or, equivalently, $\overline{A} = \Theta B \Theta^T \circ \boldsymbol{\delta}\boldsymbol{\delta}^T$ where $\circ$ is the Hadamard (componentwise) matrix product. Again, edges are generated independently from one another with the probabilities prescribed by the corresponding entries in $\overline{A}$ (we tacitly assume that $\boldsymbol{\delta}$ and B are scaled so that all entries of $\overline{A}$ lie in $[0, 1]$).

As it will be useful in following discussion, assuming that blocks consist of consecutive integers, we also write $\boldsymbol{\delta} = (\boldsymbol{\delta}^{(1)}, \dots, \boldsymbol{\delta}^{(k)})^T$ where each $\boldsymbol{\delta}^{(i)}$ is a vector of length n_i. In the dc-SBM, equation (1) is replaced by

$$\overline{A} = \begin{pmatrix} \overline{A}_{11} & \cdots & \overline{A}_{1k} \\ \vdots & & \vdots \\ \overline{A}_{k1} & \cdots & \overline{A}_{kk} \end{pmatrix}, \qquad \overline{A}_{ij} = b_{ij}\boldsymbol{\delta}^{(i)}\boldsymbol{\delta}^{(j)T} \in \mathbb{R}^{n_i\times n_j}. \tag{2}$$

In particular, the expected value of the (i, j)-entry of a random adjacency matrix in this model is $\overline{A}_{ij} = \delta_i \delta_j b_{\beta(i),\beta(j)}$, and the expected degree of node i is

$$\overline{d}_i = \sum_{j=1}^{n} \overline{A}_{ij} = \delta_i \sum_{j=1}^{n} \delta_j b_{\beta(i),\beta(j)}. \tag{3}$$

That number may be not constant within blocks. In fact, equation (3) shows that, if p and q are two nodes in the same block then their expected degrees are proportional to δ_p and δ_q. The dc-SBM includes as special cases:

- the (Extended) Planted Partition model, where $B = \alpha I + \beta \mathbf{1}\mathbf{1}^T$ [5],
- the standard SBM, when $\boldsymbol{\delta} = \mathbf{1}$ [24], and
- the Chung–Lu random graph model, when $k = 1$ and B is the 1×1 matrix $B = 1$ [7].

3 The matrix inflation operator

We borrow from [15] the definition of inflation product of two matrices. This operator is the special case of the Khatri–Rao matrix product, see e.g. [17, §12.3.3], occurring when the diagonal blocks of the second matrix are square.

Definition 1. *Let $A \in \mathbb{R}^{n\times n}$ be a matrix partitioned in $k \times k$ block form, with square diagonal blocks,*

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & & \vdots \\ A_{k1} & \cdots & A_{kk} \end{pmatrix}, \qquad A_{ij} \in \mathbb{R}^{n_i \times n_j}$$

with $n_1 + \cdots + n_k = n$, and let $B = (b_{ij}) \in \mathbb{R}^{k\times k}$. The inflation of B with respect to A *is the $n \times n$ block matrix*

$$B \odot A = \begin{pmatrix} b_{11}A_{11} & \cdots & b_{1k}A_{1k} \\ \vdots & & \vdots \\ b_{k1}A_{k1} & \cdots & b_{kk}A_{kk} \end{pmatrix}.$$

Analogously, if $\boldsymbol{w} \in \mathbb{R}^n$ is a vector partitioned into k sub-vectors,

$$\boldsymbol{w} = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_k)^T, \qquad \boldsymbol{w}_i \in \mathbb{R}^{n_i}, \tag{4}$$

and $\boldsymbol{v} = (v_1, \dots, v_k)^T \in \mathbb{R}^k$ then

$$\boldsymbol{v} \odot \boldsymbol{w} = (v_1\boldsymbol{w}_1, \dots, v_k\boldsymbol{w}_k)^T \in \mathbb{R}^n.$$

The inflation product is a linear operator with respect to each argument, and obeys the following rule: if $\boldsymbol{v}$ and $\boldsymbol{w}$ are vectors then $(\boldsymbol{v} \odot \boldsymbol{w})(\boldsymbol{v} \odot \boldsymbol{w})^T = \boldsymbol{v}\boldsymbol{v}^T \odot \boldsymbol{w}\boldsymbol{w}^T$. Note that the notation $B \odot A$ does not mention explicitly the partitioning of A on which the result depends; that partitioning should be clear from the context. In the analysis of matrices arising from a Stochastic Block Model, we refer to the block partitioning described by the membership matrix Θ, as shown in (1).

The operator $\odot$ is closely related to the Kronecker product $\otimes$; indeed, when $n_1 = \ldots = n_k = n/k$ and all blocks of A are equal to Z then $B \odot A = B \otimes Z$. On the other hand, if $k = n$ then $n_1 = \ldots = n_k = 1$ and $B \odot A = B \circ A$, the Hadamard (componentwise) product. Those relations are no longer valid if the blocks have different sizes. The special case $B \odot \mathbf{1}\mathbf{1}^T$ (with arbitrary block sizes) has been considered by M. Bolla in [1] where it is named *blown up matrix*.

The following result is a simple case of more general results shown in Lemma 4.17 and Lemma 4.19 of [15]; we include here a short proof.

Lemma 1. *Let $\boldsymbol{w} \in \mathbb{R}^n$ and let $B \in \mathbb{R}^{k \times k}$ be a symmetric matrix. Suppose that $\boldsymbol{w}$ is partitioned into k sub-vectors as in (4) with $\|\boldsymbol{w}_i\|_2 = 1$ for $i = 1, \dots, k$. Let $(\lambda, \boldsymbol{v})$ be any eigenpair of B, $B\boldsymbol{v} = \lambda \boldsymbol{v}$. Then,*

$$(B \odot \boldsymbol{w}\boldsymbol{w}^T)(\boldsymbol{v} \odot \boldsymbol{w}) = \lambda\,(\boldsymbol{v} \odot \boldsymbol{w}).$$

Furthermore, let $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_k)^T \in \mathbb{R}^n$ be any vector partitioned consistently with $\boldsymbol{w}$ and such that $\boldsymbol{w}_i^T \boldsymbol{x}_i = 0$ for $i = 1, \dots, k$. Then $(B \odot \boldsymbol{w}\boldsymbol{w}^T)\boldsymbol{x} = 0$.

Proof. Let $\boldsymbol{v} = (v_1, \dots, v_k)^T$. For $i = 1, \dots, k$ the i-th sub-vector of $(B \odot \boldsymbol{w}\boldsymbol{w}^T)(\boldsymbol{v} \odot \boldsymbol{w})$ is given by

$$\begin{aligned}\sum_{j=1}^{k} (B \odot \boldsymbol{w}\boldsymbol{w}^T)_{ij} (\boldsymbol{v} \odot \boldsymbol{w})_j &= \sum_{j=1}^{k} (b_{ij} \boldsymbol{w}_i \boldsymbol{w}_j^T) v_j \boldsymbol{w}_j \\ &= \boldsymbol{w}_i \sum_{j=1}^{k} b_{ij} (\boldsymbol{w}_j^T \boldsymbol{w}_j) v_j = (B\boldsymbol{v})_i \boldsymbol{w}_i = \lambda v_i \boldsymbol{w}_i,\end{aligned}$$

which is the i-th sub-vector of $\lambda(\boldsymbol{v} \odot \boldsymbol{w})$. Analogously, for the second part of the claim, we have

$$\big[(B \odot \boldsymbol{w}\boldsymbol{w}^T)\boldsymbol{x}\big]_i = \sum_{j=1}^{k} (B \odot \boldsymbol{w}\boldsymbol{w}^T)_{ij} \boldsymbol{x}_j = \sum_{j=1}^{k} b_{ij} \boldsymbol{w}_i (\boldsymbol{w}_j^T \boldsymbol{x}_j) = 0$$

and the proof is complete. □

Thus, in the hypotheses stated in the previous result, if $B = V \Lambda V^T$ is a spectral decomposition of B with $V = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_n]$ being an orthogonal matrix, then we have the spectral decomposition

$$B \odot \boldsymbol{w}\boldsymbol{w}^T = W \Lambda W^T, \qquad W = [\boldsymbol{v}_1 \odot \boldsymbol{w}, \dots, \boldsymbol{v}_n \odot \boldsymbol{w}].$$

In particular, the nonzero eigenvalues of $B \odot \boldsymbol{w}\boldsymbol{w}^T$ coincide with the nonzero eigenvalues of B. Remark that, if some sub-vector of $\boldsymbol{w}$ is zero then the corresponding block of rows and columns in $B \odot \boldsymbol{w}\boldsymbol{w}^T$ vanishes. This fact may reduce the rank and introduce more zero eigenvalues in $B \odot \boldsymbol{w}\boldsymbol{w}^T$. In that case the nonzero eigenvalues can be characterized as eigenvalues of a suitable submatrix of B by means of a deflation argument. In the next theorem we relax the condition on the subvectors $\boldsymbol{w}_i$ stated in the previous lemma.

Theorem 1. *In the same notations of Lemma 1, let $z_i = \|\boldsymbol{w}_i\|_2 \neq 0$ for $i = 1, \dots, k$ and let $Z = \mathrm{Diag}(z_1, \dots, z_k)$. If $\widehat{B} = V \Lambda V^T$ is a spectral decomposition of $\widehat{B} = ZBZ$ with $V = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_k]$ orthogonal then a spectral decomposition of $B \odot \boldsymbol{w}\boldsymbol{w}^T$ is*

$$B \odot \boldsymbol{w}\boldsymbol{w}^T = W \Lambda W^T, \qquad W = [Z^{-1}\boldsymbol{v}_1 \odot \boldsymbol{w}, \dots, Z^{-1}\boldsymbol{v}_k \odot \boldsymbol{w}].$$

In particular, the nonzero eigenvalues of $B \odot \boldsymbol{w}\boldsymbol{w}^T$ coincide with the nonzero eigenvalues of $\widehat{B}$.

Proof. Define $\hat{\boldsymbol{w}}_i = \boldsymbol{w}_i / z_i$. Clearly, $\|\hat{\boldsymbol{w}}_i\|_2 = 1$ and $\mathrm{Diag}(\hat{\boldsymbol{w}})\Theta Z = \mathrm{Diag}(\boldsymbol{w})\Theta$. Hence

$$\begin{aligned}\widehat{B} \odot \hat{\boldsymbol{w}}\hat{\boldsymbol{w}}^T &= \mathrm{Diag}(\hat{\boldsymbol{w}})\Theta \hat{B} \Theta^T \mathrm{Diag}(\hat{\boldsymbol{w}}) \\ &= \mathrm{Diag}(\boldsymbol{w})\Theta B \Theta^T \mathrm{Diag}(\boldsymbol{w}) = B \odot \boldsymbol{w}\boldsymbol{w}^T.\end{aligned}$$

Let $(\lambda, \boldsymbol{v})$ be any eigenpair of $\widehat{B}$. From Lemma 1 we obtain that $\boldsymbol{v} \odot \hat{\boldsymbol{w}}$ is an eigenvector of $\widehat{B}$ associated to the eigenvalue λ. For $i = 1, \dots, k$ the i-th sub-vector of $\boldsymbol{v} \odot \hat{\boldsymbol{w}}$ is

$$v_i \hat{\boldsymbol{w}}_i = \frac{v_i}{z_i} \boldsymbol{w}_i = (Z^{-1}\boldsymbol{v})_i \boldsymbol{w}_i,$$

which is the i-th sub-vector of $Z^{-1}\boldsymbol{v} \odot \boldsymbol{w}$, and the proof is complete. □

4 Analysis of the dc-SBM in expectation

In this section we consider a generic dc-SBM with n nodes and k blocks, identified by parameters $(\Theta, B, \boldsymbol{\delta})$. We generally assume that the matrix B is nonsingular. That hypothesis is commonly adopted in the literature, see e.g., [1, 8, 23, 24], as it is a fundamental requirement to ensure the identifiability of the block structure from random matrices drawn from the model.

The forthcoming result displays compact, explicit formulas for the average adjacency matrix and the average degree vector, in terms of the parameters of the model.

Lemma 2. *The expected adjacency matrix* (2) *of the dc-SBM* $(\Theta, B, \boldsymbol{\delta})$ *can be expressed as*

$$\overline{A} = B \odot \boldsymbol{\delta}\boldsymbol{\delta}^T.$$

Moreover, let $\overline{\boldsymbol{d}} = \overline{A}\,\mathbf{1}$ *be the expected degree vector and let* $\boldsymbol{t} = \Theta^T\boldsymbol{\delta}$. *Then,* $\overline{\boldsymbol{d}} = (B\boldsymbol{t}) \odot \boldsymbol{\delta}$.

Proof. The first part of the claim is a direct consequence of equation (2) and Definition 1. For $i = 1, \dots, k$ let $\boldsymbol{\delta}^{(i)}$ and $\overline{\boldsymbol{d}}^{(i)}$ be the i-th subvector of $\boldsymbol{\delta}$ and $\overline{\boldsymbol{d}}$, respectively. Noting that $\boldsymbol{t} = (t_1, \dots, t_k)^T$ is the vector with $t_i = \mathbf{1}^T\boldsymbol{\delta}^{(i)}$, from (3) we obtain

$$\overline{\boldsymbol{d}}^{(i)} = \Big(\sum_{j=1}^{n} b_{i,\beta(j)}\delta_j\Big)\boldsymbol{\delta}^{(i)} = \Big(\sum_{\ell=1}^{k} b_{i\ell}\,\mathbf{1}^T\boldsymbol{\delta}^{(\ell)}\Big)\boldsymbol{\delta}^{(i)} = (B\boldsymbol{t})_i\boldsymbol{\delta}^{(i)},$$

and we have the claim. □

The preceding lemma provides a compact form for the expected degree vector $\overline{\boldsymbol{d}}$, which clarifies the nonlinear dependence of $\overline{\boldsymbol{d}}$ on the model parameter $\boldsymbol{\delta}$. In particular, in the standard SBM it holds $\overline{\boldsymbol{d}}^{(i)} = (B\boldsymbol{t})_i\mathbf{1}$, while in the Chung–Lu model we have $\overline{\boldsymbol{d}} = (\mathbf{1}^T\boldsymbol{\delta})\boldsymbol{\delta}$.

On the other hand, in the analysis of graphs drawn from a dc-SBM one typically has an empirical knowledge of a graph and is interested in the inverse problem of computing $\boldsymbol{\delta}$, solution of $\overline{\boldsymbol{d}} = (B\boldsymbol{t}) \odot \boldsymbol{\delta}$. For $k = 2$ this problem has been addressed in [8], where the authors provide a partial solution and point out that there are examples where that equation has no solution. Owing to the identities $\overline{A} = \mathrm{Diag}(\boldsymbol{\delta})\Theta B\Theta^T\mathrm{Diag}(\boldsymbol{\delta})$ and $\overline{\boldsymbol{d}} = \overline{A}\mathbf{1}$, solving $\overline{\boldsymbol{d}} = (B\boldsymbol{t}) \odot \boldsymbol{\delta}$ for $\boldsymbol{\delta}$ boils down to the widely studied matrix theoretic problem known as matrix balancing problem, where we look for a positive diagonal scaling D of the symmetric matrix $\Theta B\Theta^T$ so that the scaled matrix $D\Theta B\Theta^T D$ has a prescribed row sum. R. Brualdi provides in [4] necessary and sufficient conditions for the existence of its solution. We outline below an immediate consequence of [4, Thm. 4].

Theorem 2. *If all entries of* B *are positive then for any nonnegative vector* $\overline{\boldsymbol{d}}$ *there exists a vector* $\boldsymbol{\delta}$ *which fulfills the identity* $\overline{\boldsymbol{d}} = (B\boldsymbol{t}) \odot \boldsymbol{\delta}$.

For completeness, let us briefly comment about related computational strategies. As mentioned before, $\boldsymbol{\delta}$ is characterized as the solution of a matrix balancing problem and thus an ad-hoc implementation of the Sinkhorn–Knopp method can be used [18]. Moreover, as $\boldsymbol{t}$ depends on $\boldsymbol{\delta}$, the equation $\overline{\boldsymbol{d}} = (B\boldsymbol{t}) \odot \boldsymbol{\delta}$ can be recast as a nonlinear fixed point problem, whose solution can be also addressed by means of a nonlinear power method, see e.g., [16]. The analysis, implementation and practical application of these methods to clustering problems for dc-SBM graphs go beyond the scope of this paper and will be the subject of a dedicated future work.

In what follows, we provide a detailed spectral analysis of the expected adjacency and modularity matrices of the dc-SBM, in view of the possible applications to spectral methods in graph clustering and community detection.

Theorem 3. *Let* $z_i = \|\boldsymbol{\delta}^{(i)}\|_2 \neq 0$ *for* $i = 1, \dots, k$ *and let* $Z = \mathrm{Diag}(z_1, \dots, z_k)$. *The nonzero eigenvalues of* $\overline{A}$ *coincide with the (nonzero) eigenvalues of the* $k \times k$ *matrix* $\widehat{B} = ZBZ$. *Furthermore, if* $\widehat{B}\boldsymbol{x} = \lambda\boldsymbol{x}$ *then* $\overline{A}(Z^{-1}\boldsymbol{x} \odot \boldsymbol{\delta}) = \lambda(Z^{-1}\boldsymbol{x} \odot \boldsymbol{\delta})$.

Proof. The claim is an immediate consequence of Lemma 2 and Theorem 1. □

Corollary 1. *In the notations of Lemma 2, let* $\boldsymbol{b} = B\boldsymbol{t}$ *and* $v = \overline{\boldsymbol{d}}^T\mathbf{1}$. *The expected Newman–Girvan modularity matrix of the dc-SBM with parameters* $(\Theta, B, \boldsymbol{\delta})$ *is*

$$\overline{M} = (B - \boldsymbol{b}\boldsymbol{b}^T/v) \odot \boldsymbol{\delta}\boldsymbol{\delta}^T.$$

The nonzero eigenvalues of $\overline{M}$ *coincide with the nonzero eigenvalues of the* $k \times k$ *matrix* $\widehat{M} = Z(B - \boldsymbol{b}\boldsymbol{b}^T/v)Z$ *where* $Z = \mathrm{Diag}(\|\boldsymbol{\delta}^{(1)}\|_2, \dots, \|\boldsymbol{\delta}^{(k)}\|_2)$. *Furthermore, let* $(\lambda, \boldsymbol{v})$ *be an eigenpair of* $\widehat{M}$ *and let* $\boldsymbol{x} = (Z^{-1}\boldsymbol{v}) \odot \boldsymbol{\delta}$. *Then* $\overline{M}\boldsymbol{x} = \lambda\boldsymbol{x}$.

Proof. By linearity of the expectation, $\overline{M} = \overline{A} - R$ for some symmetric rank-one matrix R. Since the equation $M\mathbf{1} = 0$ holds true for every Newman–Girvan modularity matrix M, we must have $R\mathbf{1} = \overline{A}\,\mathbf{1} = \overline{\boldsymbol{d}}$, hence $R = \overline{\boldsymbol{d}}\,\overline{\boldsymbol{d}}^T/(\overline{\boldsymbol{d}}^T\mathbf{1})$. Therefore,

$$\begin{aligned}\overline{M} &= B \odot \boldsymbol{\delta}\boldsymbol{\delta}^T - (\boldsymbol{b} \odot \boldsymbol{\delta})(\boldsymbol{b} \odot \boldsymbol{\delta})^T/v \\ &= B \odot \boldsymbol{\delta}\boldsymbol{\delta}^T - (\boldsymbol{b}\boldsymbol{b}^T/v) \odot \boldsymbol{\delta}\boldsymbol{\delta}^T = (B - \boldsymbol{b}\boldsymbol{b}^T/v) \odot \boldsymbol{\delta}\boldsymbol{\delta}^T,\end{aligned}$$

and we obtain the first part of the claim. Finally, the spectral properties of $\overline{M}$ mentioned above follow at once from Theorem 1. □

Remark 1. *We point out that the matrix* $B - \boldsymbol{b}\boldsymbol{b}^T/v$ *appearing in the previous corollary has a nontrivial kernel. Indeed, from* $\boldsymbol{t} = (\mathbf{1}^T\boldsymbol{\delta}^{(1)}, \dots, \mathbf{1}^T\boldsymbol{\delta}^{(k)})^T$, $\boldsymbol{b} = B\boldsymbol{t}$ *and* $\overline{\boldsymbol{d}} = (B\boldsymbol{t}) \odot \boldsymbol{\delta}$ *we get*

$$\begin{aligned}\boldsymbol{b}^T\boldsymbol{t} = \boldsymbol{t}^T B\boldsymbol{t} &= \sum_{i=1}^{k}(B\boldsymbol{t})_i\,\mathbf{1}^T\boldsymbol{\delta}^{(i)} \\ &= \sum_{i=1}^{k}\mathbf{1}^T\overline{\boldsymbol{d}}^{(i)} = \mathbf{1}^T\overline{\boldsymbol{d}} = v,\end{aligned}$$

and we conclude that $(B - \boldsymbol{b}\boldsymbol{b}^T/v)\boldsymbol{t} = \boldsymbol{b} - (\boldsymbol{b}^T\boldsymbol{t}/v)\boldsymbol{b} = 0$. *By congruence, that zero eigenvalue enters also in the matrix* $\widehat{M}$ *with corresponding eigenvector* $Z^{-1}\boldsymbol{t}$ *and, as a consequence of Corollary 1, also in the spectrum of* $\overline{M}$. *Hence, the matrix* $\overline{M}$ *for a dc-SBM with k blocks has at most* $k-1$ *nonzero eigenvalues. Furthermore, if B is nonsingular then they are exactly* $k-1$, *by the eigenvalue interlacing theorem. This fact confirms that the presence of k principal clusters in a graph drawn from stochastic block models is indicated by the presence of* $k-1$ *dominant eigenvalues in the spectrum of the modularity matrix, as discussed for instance in [11, 14].*

4.1 Normalized adjacency and modularity matrices

Let A be the adjacency matrix of a graph, and let $\boldsymbol{d} = A\mathbf{1}$ be the degree vector. Define the matrix $D = \mathrm{Diag}(\boldsymbol{d})$. Then the matrices $\mathcal{A} = D^{-1/2}AD^{-1/2}$ and $\mathcal{M} = D^{-1/2}MD^{-1/2}$ are the *normalized adjacency matrix* and *normalized modularity matrix*, respectively. These matrices are often encountered in the literature on spectral graph analysis, see e.g., [2, 7, 8, 12, 13, 23]. Among various reasons, their relevance stems from the fact that the Perron eigenvalue of $\mathcal{A}$ is equal to one, and the corresponding Perron vector is $\boldsymbol{x} = D^{1/2}\mathbf{1}$. Moreover, $\mathcal{A}$ and $\mathcal{M}$ are simultaneously diagonalizable because of the identity $\mathcal{M} = \mathcal{A} - \boldsymbol{x}\boldsymbol{x}^T/\boldsymbol{x}^T\boldsymbol{x}$, and $\mathcal{M}\boldsymbol{x} = 0$.

Due to the nonlinear dependence of both $\overline{A}$ and $\overline{\boldsymbol{d}}$ on the parameters of the model, providing an explicit formula for the expectation of the normalized versions of the adjacency and modularity matrices in a dc-SBM is not trivial. This is one of the reasons why, in many situations, the normalized version of the expected matrices is used as a replacement (see for instance [24]). For the adjacency matrix, this is defined as $\overline{\mathcal{A}} = \overline{D}^{-1/2}\overline{A}\,\overline{D}^{-1/2}$ where $\overline{D} = \mathrm{Diag}(\overline{\boldsymbol{d}})$. The matrix $\overline{\mathcal{A}}$ has various desirable properties. For instance, simple computations prove the identity

$$\overline{\mathcal{A}}\overline{\boldsymbol{x}} = \overline{\boldsymbol{x}}, \qquad \overline{\boldsymbol{x}} = \overline{D}^{1/2}\mathbf{1}\,.$$

This shows that $\overline{\mathcal{A}}$, as any normalized adjacency matrix, has unit spectral norm. This is one of the most relevant properties of a normalized adjacency matrix which, however, is not necessarily shared by the expected normalized adjacency matrix. Moreover, a direct inspection shows that

$$\overline{\mathcal{A}} = B \odot \boldsymbol{\tau}\boldsymbol{\tau}^T, \qquad \boldsymbol{\tau} = \overline{D}^{-1/2}\boldsymbol{\delta}.$$

As a result, the spectral analysis of the matrix $\overline{\mathcal{A}}$ follows from Theorem 2 by replacing $\boldsymbol{\delta}$ with $\boldsymbol{\tau}$.

Analogously, define $\overline{\mathcal{M}} = \overline{D}^{-1/2}\overline{M}\,\overline{D}^{-1/2}$. On the basis of the equation

$$\overline{\mathcal{M}} = \overline{\mathcal{A}} - \overline{\boldsymbol{x}}\,\overline{\boldsymbol{x}}^T/\overline{\boldsymbol{x}}^T\overline{\boldsymbol{x}},$$

it is straightforward to observe that $\overline{\mathcal{A}}$ and $\overline{\mathcal{M}}$ are simultaneously diagonalizable, and $\overline{\mathcal{M}}\,\overline{\boldsymbol{x}} = 0$. Hence, the spectral relationships between these two matrices mirror the ones existing between usual normalized adjacency and modularity matrices. Furthermore, in the same notations of Corollary 1, we have

$$\overline{\mathcal{M}} = (B - \boldsymbol{b}\boldsymbol{b}^T/\nu) \odot \boldsymbol{\tau}\boldsymbol{\tau}^T,$$

and the spectral analysis of $\overline{\mathcal{M}}$ is an immediate consequence of that corollary.

4.2 The collapsed graph

Let $T = \mathrm{Diag}(\boldsymbol{t})$ with $t_i = \mathbf{1}^T\boldsymbol{\delta}^{(i)}$. It is interesting to regard the matrix TBT as the adjacency matrix of a complete graph with k nodes and the weight $b_{ij}t_it_j$ placed on the edge ij (if $i = j$ then the edge is a loop). This graph can be regarded as a "collapsed" version of the expected graph in the $(\Theta, B, \boldsymbol{\delta})$-model. Indeed,

$$\Theta^T\overline{A}\Theta = \Theta^T\Delta\Theta B\Theta^T\Delta\Theta = TBT.$$

Hence, every node of the small graph corresponds to the merging of all nodes belonging to the same block of the expected graph in a single macro-node; edges are merged consistently, in the most obvious way. Consequently, the "degree vector" of the collapsed graph is

$$TBT\mathbf{1} = TB\boldsymbol{t} = T\boldsymbol{b}.$$

The i-th entry of $T\boldsymbol{b}$ is equal to $t_ib_i = (B\boldsymbol{t})_i\mathbf{1}^T\boldsymbol{\delta}^{(i)} = \mathbf{1}^T\overline{\boldsymbol{d}}^{(i)}$, which is the expected volume of the i-th block in the model; and the volume of the collapsed graph is $\mathbf{1}^TT\boldsymbol{b} = \boldsymbol{t}^T\boldsymbol{b} = \nu$, the expected volume in the model. Furthermore, simple computations show that the Newman–Girvan modularity matrix corresponding to TBT admits the equivalent formulations

$$\Theta^T\overline{M}\Theta = \Theta^T((B - \boldsymbol{b}\boldsymbol{b}^T/\nu) \odot \boldsymbol{\delta}\boldsymbol{\delta}^T)\Theta = T(B - \boldsymbol{b}\boldsymbol{b}^T/\nu)T,$$

so that the transition to the collapsed graph preserves the expected modularity of the blocks and of their unions. Obviously, the map $\overline{A} \mapsto \Theta^T\overline{A}\Theta$ does not preserve the nonzero eigenvalues of $\overline{A}$. However, from Theorem 3 and Corollary 1 we derive that the map $\overline{A} \mapsto (T^{-1}Z)^T\Theta^T\overline{A}\Theta(T^{-1}Z)$ does it, and the same is true also for the analogous transform of $\overline{M}$.

4.3 Recovering the block structure

In this section we address the problem of recovering the complete block structure induced by Θ from the spectral analysis of the expected adjacency and modularity matrices $\overline{A}$, $\overline{M}$. In what follows, we use the widespread terms *structural eigenvalues* and *structural eigenvectors* to indicate the nonzero eigenvalues of $\overline{A}$ and the corresponding eigenvectors.

As shown in Theorem 3, if B is nonsingular and $\boldsymbol{\delta}^{(i)} \neq 0$ for $i = 1, \dots, k$ then $\overline{A}$ has k structural eigenvalues. Moreover, the $n \times k$ matrix formed by the structural eigenvectors is

$$U_A = [Z^{-1}\boldsymbol{v}_1 \odot \boldsymbol{\delta}, \dots, Z^{-1}\boldsymbol{v}_k \odot \boldsymbol{\delta}]$$

where $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ are orthonormal eigenvectors of $\widehat{B} = ZBZ$. Let $V_A = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_k]$. Introduce the matrix $W = Z^{-1}V_A$ and denote its rows by $\boldsymbol{w}_1^T, \dots, \boldsymbol{w}_k^T$, so that $\boldsymbol{w}_i^T = \boldsymbol{e}_i^T W$ where $\boldsymbol{e}_i$ is the i-th column of the identity matrix of order k. Then we have the equivalent formulas

$$U_A = \begin{pmatrix} \boldsymbol{\delta}^{(1)}\boldsymbol{e}_1^T \\ \vdots \\ \boldsymbol{\delta}^{(k)}\boldsymbol{e}_k^T \end{pmatrix} Z^{-1}V_A = \begin{pmatrix} \boldsymbol{\delta}^{(1)}\boldsymbol{w}_1^T \\ \vdots \\ \boldsymbol{\delta}^{(k)}\boldsymbol{w}_k^T \end{pmatrix}.$$

Let $i = 1, \dots, n$ and $j = \beta(i)$. From the previous formula it is apparent that the i-th row of U_A is $\delta_i \boldsymbol{w}_j^T$, which is a multiple of the j-th row of V_A. In conclusion, if two nodes belong to the same block then the corresponding rows of U_A are positive multiples of each other, otherwise they are orthogonal. This fact is well known and lies at the foundations of various clustering methods based on the k-means algorithm: A sufficiently accurate knowledge of the structural eigenvectors allows to recover the assignment of nodes to blocks on the basis of the angles formed by the rows of U_A. However, the length of the i-th row of U_A is $\delta_i/z_j = \delta_i/\|\boldsymbol{\delta}^{(j)}\|_2$. If this ratio is small then that row is more sensitive to the random perturbations due to the Wigner noise. Hence, if the expected degree of i is much lower than the average degree of nodes in the same block, then that node is likely to be misclassified.

In the remaining part of this subsection we compare the situation for the matrix $\overline{A}$ described above with that for the matrix $\overline{M}$. By Corollary 1, the matrix $\widehat{M} = Z(B - \boldsymbol{b}\boldsymbol{b}^T/\nu)Z$ establishes the nonzero part of the spectrum of $\overline{M}$. Arguing as for $\overline{A}$, if V_M is an orthonormal matrix formed by the eigenvectors of $\widehat{M}$ then the corresponding eigenvectors of $\overline{M}$ are the columns of the matrix

$$U_M = \begin{pmatrix} \boldsymbol{\delta}^{(1)}\boldsymbol{e}_1^T \\ \vdots \\ \boldsymbol{\delta}^{(k)}\boldsymbol{e}_k^T \end{pmatrix} Z^{-1}V_M. \tag{5}$$

Also in the matrix U_M, rows belonging to the same block are parallel, and rows belonging to different blocks are orthogonal. However, as anticipated in Remark 1, at least one of those eigenvectors is associated to a null eigenvalue. In a concrete graph drawn from an SBM, that eigenpair is spoiled by the Wigner noise, so that it cannot be used in any spectral method, even if the structural eigenvalues are well separated from the bulk of the spectrum. Hence, we have to assess the possibility to recover the block structure of the model on the basis of the structural eigenvectors of $\overline{M}$. The question is settled in the next result.

Theorem 4. *Let the matrix B be nonsingular, and let U_r be the $n \times r$ matrix whose columns are the eigenvectors of $\overline{M}$ associated to its nonzero eigenvalues, $r = \operatorname{rank}(\overline{M})$. If i and j are two indices belonging to the same block in the block structure induced by Θ then the corresponding rows of U_r are positive multiple of each other; otherwise, they form an obtuse angle.*

Proof. If B is nonsingular then, as observed in Remark 1, the rank of $\overline{M}$ is $r = k - 1$. Without loss in generality, we can assume that U_r is formed by the first r columns of the matrix U_M in (5). Now, let V_r be formed by the first r columns of V_M and let $\boldsymbol{x}_i^T$ be the i-th row of $Z^{-1}V_r$. Then,

$$U_r = \begin{pmatrix} \boldsymbol{\delta}^{(1)}\boldsymbol{e}_1^T \\ \vdots \\ \boldsymbol{\delta}^{(k)}\boldsymbol{e}_k^T \end{pmatrix} Z^{-1}V_r = \begin{pmatrix} \boldsymbol{\delta}^{(1)}\boldsymbol{x}_1^T \\ \vdots \\ \boldsymbol{\delta}^{(k)}\boldsymbol{x}_k^T \end{pmatrix}.$$

Hence, the i-th row of U_r is $\boldsymbol{u}_i^T = \delta_i \boldsymbol{x}_{\beta(i)}^T$. As a consequence, for any two indices $i, j = 1, \dots, n$ the angle between $\boldsymbol{u}_i^T$ and $\boldsymbol{u}_j^T$ is the same as the one between $\boldsymbol{x}_{\beta(i)}^T$ and $\boldsymbol{x}_{\beta(j)}^T$. In particular, if $\beta(i) = \beta(j)$ then $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ are positive multiple of each other. It remains to analyze the case where $\beta(i) \neq \beta(j)$.

The matrix V_r is made by the first $k-1$ columns of the matrix V_M, the missing column being a normalized eigenvector associated to the zero eigenvalue of $\widehat{M}$. According to Remark 1, that eigenvector can be expressed as $\hat{\boldsymbol{v}} = (\hat{v}_1, \dots, \hat{v}_k)^T = Z^{-1}\boldsymbol{t}/\|Z^{-1}\boldsymbol{t}\|_2$. Recall that, by initial assumptions, $t_i = \mathbf{1}^T\boldsymbol{\delta}^{(i)} > 0$ and $z_i = \|\boldsymbol{\delta}^{(i)}\|_2 > 0$, hence $\hat{v}_i > 0$. Since V_M is orthogonal,

$$[\boldsymbol{x}_1, \dots, \boldsymbol{x}_k]^T[\boldsymbol{x}_1, \dots, \boldsymbol{x}_k] = Z^{-1}V_r^T V_r Z^{-1} = Z^{-1}(I - \hat{\boldsymbol{v}}\hat{\boldsymbol{v}}^T)Z^{-1},$$

so for $i, j = 1, \dots, k$ we have

$$\boldsymbol{x}_i^T\boldsymbol{x}_j = (\boldsymbol{e}_i^T\boldsymbol{e}_j - \hat{v}_i\hat{v}_j)/(z_i z_j).$$

Note that $\|\boldsymbol{x}_i\|_2 > 0$ since $\hat{v}_i^2 < 1$. Thus if $i \neq j$ the cosine of the angle between $\boldsymbol{x}_i^T$ and $\boldsymbol{x}_j^T$ is

$$\frac{\boldsymbol{x}_i^T\boldsymbol{x}_j}{\|\boldsymbol{x}_i\|_2\|\boldsymbol{x}_j\|_2} = \frac{-\hat{v}_i\hat{v}_j}{(1-\hat{v}_i^2)^{1/2}(1-\hat{v}_j^2)^{1/2}} < 0,$$

and the proof is complete. □

5 Asymptotic results

Theoretical studies on the statistical effectiveness of spectral methods for data clustering and community detection need to consider sequences of block models with growing size but sharing a common underlying structure, see e.g., [5, 8, 23, 24], in order to draw justifiable inferences about the latent block structure from the estimated clustering in large network. Spectral methods strongly rely on the magnitude of the structural eigenvalues of the adjacency or the modularity matrix. Indeed, as mentioned in the Introduction, when the graph follows a dc-SBM, these matrices differ from their expectation by a random Wigner noise and thus it is important to analyze the behavior of the structural eigenvalues with respect to the network size and compare it to the magnitude of the noise perturbation.

Suppose we have a sequence of dc-SBMs, $\{(\Theta_n, B, \boldsymbol{\delta}_n)\}$ with growing n and fixed B. Assuming rather weak hypotheses on the model parameters, the next result describes the asymptotic behavior of such structural eigenvalues for the sequence of expected adjacency matrices $\overline{A}_n$ and, consequently, also of the corresponding expected modularity matrices.

Theorem 5. *Let $\{(\Theta_n, B, \boldsymbol{\delta}_n)\}$ be a sequence of dc-SBMs, with growing n and fixed B and k. Let $\overline{A}_n = B \odot \boldsymbol{\delta}_n\boldsymbol{\delta}_n^T$ and $\boldsymbol{\delta}_n = (\boldsymbol{\delta}_n^{(1)}, \dots, \boldsymbol{\delta}_n^{(k)})^T$ with the block partitioning induced by Θ_n. Suppose that there exist nonnegative constants $\ell_1, \dots, \ell_k$ and a function $f : \mathbb{N} \mapsto \mathbb{R}$ such that*

$$\lim_{n\to\infty} \frac{\|\boldsymbol{\delta}_n^{(i)}\|_2}{f(n)} = \ell_i, \qquad i = 1, \dots, k.$$

Moreover, let $L = \mathrm{Diag}(\ell_1, \dots, \ell_k)$ and let $\lambda_1^\star, \dots, \lambda_k^\star$ be the eigenvalues of the matrix LBL. Then the nonzero eigenvalues $\lambda_1^{(n)}, \dots, \lambda_k^{(n)}$ of $\overline{A}_n$ can be numbered so that

$$\lim_{n\to\infty} \frac{\lambda_i^{(n)}}{f(n)^2} = \lambda_i^\star, \qquad i = 1, \dots, k,$$

Proof. By Theorem 3 the nonzero eigenvalues of $\overline{A}_n$ coincide with the nonzero eigenvalues of L_nBL_n where

$$L_n = \mathrm{Diag}(\|\boldsymbol{\delta}_n^{(1)}\|_2, \dots, \|\boldsymbol{\delta}_n^{(k)}\|_2).$$

Since $\lim_{n\to\infty} L_n/f(n) = L$ by hypothesis, the matrix sequence $\{L_nBL_n/f(n)^2\}$ converges to the matrix LBL, and the claim follows by continuity of the eigenvalues with respect to matrix entries. □

As a simple application, we rederive hereafter a known result on the $\mathcal{O}(n)$ behavior of the structural eigenvalues of a sequence of stochastic block models.

Corollary 2. *Let $\{(\Theta_n, B)\}$ be a sequence of SBMs with growing n and fixed B and k. If there exist nonnegative constants $c_1, \dots, c_k$ such that $\lim_{n\to\infty} n_i/n = c_i$ then the structural eigenvalues $\lambda_1^{(n)}, \dots, \lambda_k^{(n)}$ of the expected adjacency matrix $\overline{A}_n$ can be numbered so that*

$$\lim_{n\to\infty} \frac{\lambda_i^{(n)}}{n} = \lambda_i^\star, \qquad i = 1, \dots, k,$$

where $\lambda_1^\star, \dots, \lambda_k^\star$ are the eigenvalues of the matrix CBC with $C = \mathrm{Diag}(c_1, \dots, c_k)$.

Proof. It suffices to apply Theorem 5 with $\boldsymbol{\delta}_n = \mathbf{1}$ and $f(n) = \sqrt{n}$. □

We stress that the hypotheses of Theorem 5 allow great generality and can be fulfilled by more general sequences of stochastic block models, for instance, when the entries of $\boldsymbol{\delta}_n$ behave as the sampling of a suitably scaled, nonnegative, integrable function. Notably, the latter circumstance includes certain graph sequences having an asymptotic degree distribution, a recurring situation in many time-evolving, real world networks [10]. We include a simple example here below, which exploits the following concept of equally distributed sequences, borrowed from [27]: Let $\mathcal{U} = \{\{u_{in}\}_{i=1}^n\}_{n=1}^\infty$ and $\mathcal{V} = \{\{v_{in}\}_{i=1}^n\}_{n=1}^\infty$ be two doubly indexed sequences of real numbers. The sequences $\mathcal{U}, \mathcal{V}$ are *equally distributed* if for any continuous function F with bounded support,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n F(u_{in}) - F(v_{in}) = 0.$$

Example 1. *For the sake of simplicity, suppose that in every block of the partitioning induced by Θ_n there are m indices, $m = n/k$, and $\boldsymbol{\delta}_n^{(1)} = \dots = \boldsymbol{\delta}_n^{(k)} = \gamma_n$ for some vector $\gamma_n \in \mathbb{R}^m$. Let $\varphi : (0, 1] \mapsto \mathbb{R}$ be a nonnegative, continuous, square integrable function such that the entries in γ_n are a permutation of the set $\{\sigma(n)\varphi(j/m)\}_{j=1,\dots,m}$. Here, $\sigma : \mathbb{N} \mapsto \mathbb{R}$ is an arbitrary scaling function. Then, according to Lemma 2, the expected degree vector of the nodes belonging to the i-th block is given by*

$$\overline{\boldsymbol{d}}_n^{(i)} = \alpha_n \beta_i \gamma_n, \qquad \alpha_n = \sigma(n) \sum_{i=1}^m \varphi(i/m), \qquad \beta_i = \sum_{j=1}^k b_{ij}.$$

The quantity $\frac{1}{m}\sum_{i=1}^m \varphi(i/m)$ can be regarded as a quadrature formula for the integral $\int_0^1 \varphi(t)\,\mathrm{d}t$, so for sufficiently large n we can use the approximation $\alpha_n \approx m\sigma(n)\|\varphi\|_{L^1}$. Hence, for $i = 1, \dots, k$ and $c(n) = 1/(m\sigma(n)^2\|\varphi\|_{L^1})$, the set of all entries in the sequence $\{c(n)\overline{\boldsymbol{d}}_n^{(i)}\}_{m=1}^\infty$ is equally distributed with the numbers

$$\beta_i \varphi(j/m), \qquad j = 1, \dots, m, \qquad m = 1, 2, \dots$$

Note that, if $\varphi(x) = x^{-1/K}$ then that sequence follows a power law with exponent $K + 1$ [7, 10].

Furthermore, treating the sum $\frac{1}{m}\sum_{i=1}^m \varphi(i/m)^2$ as a quadrature formula for $\int_0^1 \varphi(t)^2\,\mathrm{d}t$, one obtains $\|\gamma_n\|_2 \approx \sqrt{m}\sigma(n)\|\varphi\|_{L^2}$. Hence, the hypotheses of Theorem 5 are fulfilled with $f(n) = \sqrt{n/k}\,\sigma(n)$ and $\ell_1 = \dots = \ell_k = \|\varphi\|_{L^2}$. Consequently, the asymptotic behavior of the structural eigenvalues of $\overline{A}_n$ is $\mathcal{O}(n\sigma(n)^2)$.

Acknowledgement: DF acknowledges the support received by INDAM-GNCS, Italy, for his research. The work of FT has been funded by the Marie Skłodowska-Curie individual fellowship MAGNET no. 744014.

References

[1] M. Bolla, Recognizing linear structure in noisy matrices. *Linear Algebra Appl.*, 402 (2005), 228–244.
[2] 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.
[3] M. Bolla, B. Bullins, S. Chaturapruek, S. Chen, and K. Friedl. Spectral properties of modularity matrices. *Linear Algebra Appl.*, 473 (2015), 359–376.

[4] R. A. Brualdi, The DAD theorem for arbitrary row sums. *Proc. Amer. Math. Soc.*, **45** (1974), 189–194.
[5] K. Chaudhuri, F. Chung and A. Tsiatas, Spectral Clustering of Graphs with General Degrees in the Extended Planted Partition Model, *Conference on Learning Theory (COLT)*, *J. Mach. Learn. Res.*, **23** (2012), 35:1–35:23.
[6] L. Chen, Q. Yu and B. Chen, Anti-modularity and anti-community detecting in complex networks. *J. Inf. Sci.*, **275** (2014), 293–313.
[7] F. Chung and L. Lu, *Complex Graphs and Networks*, volume 107 of CBMS Regional Conf. Ser. in Math. AMS, 2006.
[8] A. Coja-Oghlan and A. Lanka, Finding planted partitions in random graphs with general degree distributions. *SIAM J. Discrete Math.*, **23** (2009/10), 1682–1714.
[9] P. Doreian, V. Batagelj and A. Ferligoj, *Generalized blockmodeling*, volume 25 of Structural Analysis in the Social Sciences. CUP, 2005.
[10] E. Estrada, *The structure of Complex Networks*. Oxford University Press, 2012.
[11] D. Fasino and F. Tudisco, A modularity based spectral method for simultaneous detection of communities and anti-communities. *Linear Algebra Appl.*, **542** (2018), 605–623.
[12] D. Fasino and F. Tudisco, Generalized modularity matrices. *Linear Algebra Appl.*, **502** (2016), 327–345.
[13] D. Fasino and F. Tudisco, Modularity bounds for clusters located by leading eigenvectors of the normalized modularity matrix. *J. Math. Inequal.*, **11** (2017), 701–714.
[14] D. Fasino and F. Tudisco, An algebraic analysis of the graph modularity. *SIAM J. Matrix Anal. Appl.*, 35 (2014), 997–1018.
[15] S. Friedland, D. Hershkowitz and H. Schneider, Matrices whose powers are M-matrices or Z-matrices. *Trans. Amer. Math. Soc.*, **300** (1987), 343–366.
[16] A. Gautier, F. Tudisco and M. Hein, The Perron–Frobenius theorem for multi-homogeneous mappings. *arXiv:1801.05034*, (2018)
[17] G. H. Golub and C. F. Van Loan, *Matrix computations*. Johns Hopkins University Press, fourth edition, 2013.
[18] P. A. Knight, The Sinkhorn–Knopp algorithm: convergence and applications. *SIAM J. Matrix Anal. Appl.* **30** (2008), 261–275.
[19] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. De Luca and S. Albayrak, Spectral analysis of signed graphs for clustering, prediction and visualization. *Proc. SIAM Int. Conf. Data Min.* (2010), 559–570
[20] 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.
[21] P. Mercado, F. Tudisco and M. Hein, Clustering signed networks with the geometric mean of Laplacians, *Adv. Neural Inf. Process. Syst.*, **29** (2016), 4421–4429.
[22] M. E. J. Newman and M. Girvan, Finding and evaluating community structure in networks. *Phys. Rev. E*, **69** (2004), 026113.
[23] T. Qin and K. Rohe, Regularized Spectral Clustering under the Degree-Corrected Stochastic Blockmodel. *Adv. Neural Inf. Process. Syst.*, **26** (2013), 3120–3128.
[24] K. Rohe, S. Chatterjee, B. Yu, Spectral clustering and the high-dimensional stochastic blockmodel. *Ann. Statist.* **39** (2011), no. 4, 1878–1915.
[25] A. Saade, F. Krzakala and L. Zdeborová, Spectral Clustering of Graphs with the Bethe Hessian. *Adv. Neural Inf. Process. Syst.*, **27** (2014), 406–414.
[26] S. E. Schaeffer, Graph clustering. *Computer Science Review*, **1** (2007), 27–64.
[27] E. E. Tyrtyshnikov, A unifying approach to some old and new theorems on distribution and clustering. *Linear Algebra Appl.*, **232** (1996), 1–43.
[28] U. Von Luxburg, A tutorial on spectral clustering. *Stat. Comput.*, **17** (2007), 395–416.