Spectral Partitioning of Random Graphs with Given Expected Degrees

Amin Coja-Oghlan[1], Andreas Goerdt[2] and André Lanka[2]

[1] Humboldt Universität zu Berlin, Institut für Informatik
Unter den Linden 6, 10099 Berlin, Germany
coja@informatik.hu-berlin.de

[2] Fakultät für Informatik, Technische Universität Chemnitz
Straße der Nationen 62, 09107 Chemnitz, Germany
{goerdt, lanka}@informatik.tu-chemnitz.de

Abstract. It is a well established fact, that – in the case of classical random graphs like variants of $G_{n,p}$ or random regular graphs – spectral methods yield efficient algorithms for clustering (e. g. colouring or bisection) problems. The theory of large networks emerging recently provides convincing evidence that such networks, albeit looking random in some sense, cannot sensibly be described by classical random graphs. A variety of new types of random graphs have been introduced. One of these types is characterized by the fact that we have a fixed expected degree sequence, that is for each vertex its expected degree is given.

Recent theoretical work confirms that spectral methods can be successfully applied to clustering problems for such random graphs, too – provided that the expected degrees are not too small, in fact $\geq \log^6 n$. In this case however the degree of *each* vertex is concentrated about its expectation. We show how to remove this restriction and apply spectral methods when the expected degrees are bounded below just by a suitable constant.

Our results rely on the observation that techniques developed for the classical sparse $G_{n,p}$ random graph (that is $p = c/n$) can be transferred to the present situation, provided we consider a suitably normalized adjacency matrix: We divide each entry of the adjacency matrix by the product of the expected degrees of the incident vertices. Given the host of spectral techniques developed for $G_{n,p}$ this observation should be of independent interest.

1 Introduction

For definiteness we specify the model of random graphs to be considered first. This model is very similar to the one considered in [11]. For further motivation see Subsection 1.2.

1.1 The model

We consider random graphs with planted partition and given expected degree sequence which are generated as follows. Let $V = \{1, \ldots, n\}$ be the set of nodes.

We fix some symmetric $k \times k$-matrix $D = (d_{ij})$ with non-negative constants as entries. Then we assign some weight $w_u > 0$ to each node $u \in V$. We let $\overline{w} = \sum_{u \in V} w_u/n$ be the arithmetic mean of the w_u's and often use $\overline{w} \cdot n = \sum_{u \in V} w_u$.

To construct the random graph $G = (V, E)$, we partition V into k disjoint subsets $V_1, \ldots, V_k$ each of size $\geq \delta n$ for some arbitrarily small but constant $\delta > 0$. The way V is split into $V_1, \ldots, V_k$ is arbitrary. We call $V_1, \ldots, V_k$ the *planted partition*. For $u \in V$ we let $\psi(u)$ denote the number of the subset u belongs to, that is $u \in V_{\psi(u)}$. We insert each edge $\{u, v\}$ independently with probability

$$d_{\psi(u),\psi(v)} \cdot \frac{w_u \cdot w_v}{\overline{w} \cdot n}.$$

Of course the parameters should be chosen such that each probability is bounded above by 1. (It has some mild technical advantages to allow for loops as we do. A loop-edge counts as 1 to the vertex-degree.) Note, the model from [11] allows for directed edges, whereas we restrict attention to undirected graphs.

Depending on the matrix D, we can model a variety of random instances of clustering problems. For example we can generate 3-colourable graphs, then $k = 3$, the V_i are the colour classes, $d_{ii} = 0$ and $d_{ij} > 0$ for $i \neq j$. Further possibilities are graphs having a small bisection, in which case the V_i are the two sides of the bisection, or graphs with subsets of vertices which are very dense or sparse... The algorithmic problem is to efficiently reconstruct the V_i (or large parts thereof) given such a random G. Note, when all w_u are the same, we get the standard random graph $G_{n,p}$ with planted partition, where $p = w_u/n$.

We denote the degree of vertex u by d_u. The expected degree of vertex u is denoted by $w'_u = \mathbf{E}[d_u]$. Then

$$w'_u = \frac{w_u}{\overline{w} \cdot n} \cdot \sum_{v \in V} w_v \cdot d_{\psi(u),\psi(v)}.$$

We let $\overline{w}' = \sum_{u \in V} w'_u/n$ be the arithmetic mean of the expected degrees w'_u. In order for our algorithm to work properly we impose the following restrictions on the model's parameters:

1. The matrix D has full rank.
2. We have $w_u \geq \varepsilon \cdot \overline{w}$ for all u, where ε is some arbitrarily small constant.
3. We have $\overline{w} \geq d$, where $d = d(\varepsilon, D, \delta)$ is a sufficiently large constant.

Our asymptotics is such that n gets large, while D, ε, δ (and therefore k and d) are fixed. Constants behind O and Ω are positive, whereas $o(1)$ even can be negative. On the other hand the weights w_u can be picked arbitrarily subject to our restrictions (in particular depending on n) and the subsets V_i with $|V_i| \geq \delta n$ are arbitrary, too.

Note, that the expected degree w'_u depends on *all* w_v's, *all* sets V_i and the matrix D. However we observe that for all u from a fixed subset V_i the quotient w_u/w'_u is constant. The following lemma collects basic properties of the expected degree. Its proof is based on simple calculations we present in Subsection 2.2.

Lemma 1.

1. *Let* u_1, u_2 *be two vertices belonging to the same set of the planted partition. Then* $w_{u_1}/w'_{u_1} = w_{u_2}/w'_{u_2}$.
2. *There exists some (large) constant* $C = C(D, \varepsilon, \delta)$ *such that for all* $u \in V$ $1/C \leq w'_u/w_u \leq C$.
3. *The expected average degree of* G $\overline{w}' = \sum_{u \in V} w'_u/n = \Theta(\overline{w})$.

Since w_u/w'_u is the same for all $u \in V_i$, we abbreviate

$$W_i = w_u/w'_u = \Theta(1), \qquad \text{moreover} \qquad W = \overline{w}/\overline{w}' = \Theta(1). \tag{1}$$

This, in particular $w'_u = \Theta(w_u)$, shows the extent to which we consider graphs with given expected degree sequence. Note that depending on the weights w_u the restrictions 2. and 3. above allow w'_u among others to be constant, independent of n.

Note, that our model allows weights following a heavy-tailed degree distribution with constant average degree such as power laws. That is the number of weights w_u is proportional to $n \cdot {w_u}^{-\beta}$ for some constant β. The degree sequence of various social and biological networks follow a power-law with $2 < \beta < 3$. For more information we refer to the papers cited in [6]. For $\beta > 2$ we have that the average weight $\overline{w}$ is constant and a lot of weights $\gg \overline{w}$, as we allow in our model.

1.2 Motivation and related literature

The analysis of large real life networks, like the internet graph, social or bibliographical networks is one of the current topics not only of Computer Science. Clearly it is important to obtain efficient algorithms adapted to the characteristics of these networks. One particular problem of interest is the problem of detecting some kind of clusters, that is subsets of vertices having extraordinarily many or few edges. Such clusters are supposed to mirror some kind of relationship among its members (= vertices of the network). Heuristics based on the eigenvalues and eigenvectors of the adjacency matrix of the network provide one of the most flexible approaches to clustering problems applied in practice. See for example [17] or the review [23] or [21]. Note that the eigenvalues and eigenvectors of *symmetric* real valued matrices, first are real valued and second can be approximated efficiently to arbitrary precision.

The relationship between spectral properties of the adjacency matrix of a graph on the one hand and clustering properties of the graph itself on the other hand is well established. Usually this relationship is based on some separation between the (absolute) values of the largest eigenvalues and the remaining eigenvalues. It has a long tradition of being exploited in practice, among others for numerical calculations. However, it is in general not easy to obtain convincing proofs certifying the quality of spectral methods in these cases, see [26] for a notable exception.

Theoretically convincing analyses of this phenomenon have been conducted in the area of random graphs. This leads to provably efficient algorithms for clustering problems in situations where purely combinatorial algorithms do not seem to work, just to cite some examples [2], [3], or [4], or the recent [22] and subsequent work such as [16]. In particular [3] has lead to further results [12], [13]. The reason for this may be that [3] is based on a rather flexible approach to obtain spectral information about random graphs [14]: Spectral information directly follows from clustering properties known to be typically present in a random graph by (inefficient) counting arguments. We apply this technique here, too.

In a recent paper [11] Dasgupta et al. extend the techniques originally developed for $G_{n,p}$ with planted partition to random graphs with given expected degrees. Such random graphs may have many vertices whose degree deviates considerably from the average degree rendering them essentially different from $G_{n,p}$. In [24] it is shown that the largest eigenvalues of a random graph with power law degree distribution are proportional to the square root of the largest degrees. Therefore the eigenvalues and the corresponding eigenvectors can hardly reveal any non-local information about the graph. Dasgupta, Hopcroft and McSherry resolve this problem by considering a suitably normalized adjacency matrix similar to the Laplacian [5]. They can retrieve the planted partition in a model similar to ours as long as the expected in-degree and out-degree of each vertex is $\geq \log^6 n$. We show that our different normalization works even when the expected degree is bounded below by a constant. This solves an open question mentioned in Section 3 of [11]. See Subsection 2.3 for an explanation of our normalization.

1.3 Techniques and result

We consider the following algorithm to reconstruct the V_i for random graphs as generated by our model. Only for technical simplicity we restrict our attention to $k = 2$, that is our partition consists only of V_1, V_2. It poses no substantial difficulties to extend the algorithm to arbitrary, yet constant k: Instead of the two eigenvectors s_1, s_2 we use k eigenvectors $s_1, \ldots, s_k$. We discuss the values possible for the C_1 used in the algorithm further below.

Algorithm 2.

Input: The adjacency matrix $A = (a_{uv})$ of some graph $G = (V, E)$ generated in our model and the expected degree sequence $w'_1, \ldots, w'_n$.

Output: A partition V'_1, V'_2 of V.

1. Calculate the expected average degree, $\overline{w}' = \sum_{u=1}^{n} w'_u / n$.
2. Construct $M = (m_{uv})$ with $m_{uv} = \overline{w}'^2 \cdot a_{uv} / (w'_u \cdot w'_v)$.
3. Let $U = \{u \in V : \sum_{v=1}^{n} m_{uv} \leq C_1 \cdot \overline{w}'\}$.
4. Construct M^* from M by deleting all entries m_{uv} with $u \notin U$ or $v \notin U$.
5. Let s_1, s_2 be the eigenvectors of M^* belonging to the two largest eigenvalues with respect to the absolute value. Scale s_i such that $\|s_i\| = \sqrt{n}$.
6. If neither s_1 nor s_2 has the property

There are $c_1, c_2 \in \mathbb{R}$ with $|c_1 - c_2| > 1/4$ such that more than $n \cdot \sqrt{C_1/\overline{w}'}$ vertices $v \in U$ have $|s_i(v) - c_1| \leq 1/32$ and more than $n \cdot \sqrt{C_1/\overline{w}'}$ vertices have $|s_i(v) - c_2| \leq 1/32$.

set $V_1 = V$ and $V_2 = \emptyset$. Otherwise, let s be such an eigenvector. Let V_1' be the vertices whose corresponding entries in s are closer to c_1 than to c_2 and set $V_2' = V \setminus V_1'$.

Some remarks are in order. First observe that the algorithm besides the graph needs the expected degree sequence as additional information. Note that the algorithm of [11] even gets the weights w_u themselves. In case of dense graphs as in [11] w.h.p. for all $u \in V$ the actual degree d_u is asymptotically equal to its expectation w_u', that is $d_u = w_u' \cdot (1+o(1))$. This can be shown with Chernoff-like bounds as Theorem 2.8 in [18]. So, the expected degree w_u' can be approximated by the actual degree d_u for each $u \in V$. The algorithm in [11] gets the w_u and implicitly the w_u'. In contrast our algorithm only needs the w_u'. We point out, our algorithm can also use the weights w_u instead of w_u'. The analysis has to be adapted, but becomes somewhat simpler.

Of course, a natural idea is to divide the entries by the *actual* degrees rather than the expected degrees, in order to remove the requirement that the w_u' are given as additional input. It turns out that this approach can be carried out successfully, i.e., the resulting matrix is suitable to recover the planted partition as well. The analysis is technically significantly more involved, and will be given in a subsequent paper.

The novel idea is our normalization of the adjacency matrix performed in Step 2. In Subsection 2.3 we show that this normalization yields a situation formally similar to the situation of $G_{n,p}$-random graphs with planted partition and the adjacency matrix as already considered in [3] and [22]. Step 4. has the analogous effect on the spectrum of M as has the deletion of high degree vertices in the case of sparse $G_{n,p}$–graphs on the spectrum of the adjacency matrix [12]. An analogous step is also present in [3].

The value of C_1 in the algorithm can be choosen almost arbitrarily as long as it is not too large and not too small. The lower bound

$$C_1 \geq \frac{5}{W} \cdot \max_{i,j}\{d_{ij} \cdot W_i \cdot W_j\} = \Theta(1). \tag{2}$$

ensures that almost surely only a small number of vertices is deleted in Step 4., namely $|V \setminus U| \leq \exp(-\Omega(\overline{w}')) \cdot n$. We will prove this fact as inequality (17) in Subsection 4.1. Due to the lack of information Algorithm 2 is unable to calculate the bound in (2). In order to fulfill (2) one can choose $C_1 = \ln \overline{w}'$ (or some other slow-growing function of $\overline{w}'$), as we can assume that $\overline{w}' = \overline{w}/W \geq d/W$ is large enough, see restriction 3. of our model. On the other hand, C_1 has to be substantially smaller than $\overline{w}'$, e.g. $C_1 = \overline{w}'/\ln \overline{w}'$. Otherwise, the spectral gap of M^* would be too small. Lemma 6 shows this connection.

Note, the concrete values for c_1, c_2 in Step 6. depend on the model parameters and are unknown to the algorithm. So it has to find c_1 and c_2 by analyzing s_i. We point out, it suffices to have $n \cdot \sqrt{C_1/\overline{w}'}$ coordinates near to c_1 resp. c_2.

Having significantly more than $n \cdot \sqrt{C_1/\overline{w}'}$ coordinates near c_i yields that w.h.p. (up to $O(n \cdot C_1/\overline{w}')$) all coordinates have $|s(v) - c_i| \leq 1/32$.

Theorem 3. *Let D, ε, δ as specified above. and G be some graph generated by the model. With probability $1 - o(1)$ with respect to G Algorithm 2 produces a partition which differs from the planted partition V_1, V_2 only in $O(C_1 \cdot n/\overline{w}')$ vertices.*

Note that the number of vertices not classified correctly is $O(C_1 \cdot n/\overline{w}')$ and thus decreases in $\overline{w}'$ as long as $C_1 \ll \overline{w}'$. We present the proof of Theorem 3 in Section 3. The following section contains basic considerations.

2 Basic facts

2.1 Notation

We often use the following notation.

1. $\|\cdot\|$ denotes the l_2-norm of a vector or matrix.
2. The transpose of a matrix or vector M is written as M^t.
3. We abbreviate $(1, \ldots, 1)^t$ by $\mathbf{1}$.
4. The x-th component of some vector v is denoted by $v(x)$.
5. For $X \subseteq \mathbb{N}$ and a vector v the vector $v_{|X}$ is obtained from v by setting $v(x) := 0$ if $x \notin X$.
6. For some matrix M and $X, Y \subseteq \mathbb{N}$, the submatrix induced by X and Y is referred as $M_{X \times Y}$. $M_{X \times Y}$ is obtained from M by *deleting* all rows x with $x \notin X$ and all columns y with $y \notin Y$. For some vector v v_X is defined analogously. Note the difference to $v_{|X}$.
7. For a matrix $M = (m_{uv})$ we define

$$s_M(X, Y) = \sum_{\substack{x \in X \\ y \in Y}} m_{xy}.$$

We omit the parenthesis in cases like $s_M(\{u\}, Y)$ and simply write $s_M(u, Y)$.

2.2 Proof of Lemma 1

Without loss of generality we show the first and the second item for u_1, $u_2 \in V_1$, the first set of our partition. Let $u \in V_1$ be arbitrary. We have that

$$\mathbf{E}\left[d_u\right] = w'_u = \sum_{i=1}^{k} \sum_{v \in V_i} d_{1i} \cdot \frac{w_u \cdot w_v}{\overline{w} \cdot n}.$$

Dividing this by $w_u > 0$ we get

$$\frac{w'_u}{w_u} = \sum_{i=1}^{k} \sum_{v \in V_i} d_{1i} \cdot \frac{w_v}{\overline{w} \cdot n}\,, \tag{3}$$

which does not depend on $u \in V_1$. This shows the first item.

We come to the second item. As $w_v \geq \varepsilon \cdot \overline{w}$ for all $v \in V$ and (3) we have

$$\frac{w'_u}{w_u} \geq \sum_{i=1}^{k} \sum_{v \in V_i} d_{1i} \cdot \frac{\varepsilon \cdot \overline{w}}{\overline{w} \cdot n} \geq \sum_{i=1}^{k} d_{1i} \cdot |V_i| \cdot \frac{\varepsilon}{n} \geq \delta \cdot \varepsilon \cdot \sum_{i=1}^{k} d_{1i} \, .$$

Since all d_{ij} are non-negative, we have $\sum_{i=1}^{k} d_{1i} \geq 0$. Equality can be ruled out. Otherwise, D would contain a 0-row and had a rank $< k$. So, $\sum_{i=1}^{k} d_{1i}$ is bounded away from 0 by some constant and

$$w'_u / w_u \geq 1/C$$

for some large positive constant C depending on D, ε and δ but neither on $w_1, \ldots, w_n$ nor n. Using (3) again, we get

$$\frac{w'_u}{w_u} \leq \max_j \{d_{1j}\} \cdot \sum_{i=1}^{k} \sum_{v \in V_i} \frac{w_v}{\overline{w} \cdot n} = \max_j \{d_{1j}\} \leq C$$

for $C = C(D)$ large enough. The third item is an immediate consequence of the second one. □

2.3 The idea of our normalization

In case of random graphs with planted partition based on the $G_{n,p}$-model the adjacency matrix A can be used to detect (at least) large parts of the partition. The partition can be reconstructed using A's eigenvectors. The techniques are introduced in [3] for the special case of a planted 3-colouring. In the most interesting sparse case, that is $p \cdot n = O(1)$ the adjacency matrix needs to be modified in so far that vertices with large degrees are deleted. This is necessary as otherwise the largest eigenvalues of A are simply the square roots of the highest degrees [19]. W.h.p. more than $\sqrt{n}$ vertices have a degree of at least $\log\log n$, leading to more than $\sqrt{n}$ eigenvalues $\geq \sqrt{\log\log n}$. If $u_1, \ldots, u_l$ are these vertices, the eigenvectors to the largest eigenvalues essentially belong to the space spanned by $\mathbf{1}_{|\{u_1\}}, \ldots, \mathbf{1}_{|\{u_l\}}$. That makes them useless for detecting a planted partition.

This aforementioned deletion trick cannot be used for our model, because in the case of a degree distribution with a heavy tail significant parts of the graph may just be ignored in this way.

So it seems to be necessary to transform the adjacency matrix to another matrix, whose spectral properties reflect global structures as the planted partition. An approach used often is the normalized Laplacian matrix $\mathcal{L} = I - L$ with $L = (l_{uv})$ with

$$l_{uv} = \begin{cases} 1/\sqrt{d_u \cdot d_v} & \{u, v\} \in E \\ 0 & \text{otherwise} \end{cases} ,$$

where d_u, d_v is the degree of u resp. v. For more information see Chung's book [5].

A normalization of the adjacency matrix similar to the Laplacian is used in [11]. The authors divide each entry of the adjacency matrix by $\sqrt{w_u \cdot w_v}$, where w_u and w_v are the weights of the incident vertices u and v. Note, Dasgupta et al. use the weights for the normalization, neither the actual degrees nor the expected degrees. Their normalization implies that the variances of the entries inside the submatrices induced by $V_i \times V_j$ are asymptotically equal.

In contrast, we use another normalization whose analysis is somewhat easier, especially in the sparse case: Each entry of the adjacency matrix A is divided by the product of the *expected* degrees $w'_u \cdot w'_v$ of the incident vertices u and v (and multiplied with $\overline{w}'^2$). Let $M = (m_{uv})$ be the normalized matrix, in our case the expectation of m_{uv} with $u \in V_i$ and $v \in V_j$ is

$$\begin{aligned}\mathbf{E}\left[m_{uv}\right] &= \frac{\overline{w}'^2}{w'_u \cdot w'_v} \cdot \mathbf{Pr}\left[\{u,v\} \in E\right] + 0 \cdot \mathbf{Pr}\left[\{u,v\} \notin E\right] \\ &= \frac{\overline{w}'^2}{w'_u \cdot w'_v} \cdot d_{ij} \cdot \frac{w_u \cdot w_v}{\overline{w} \cdot n} \overset{(1)}{=} d_{ij} \cdot \frac{W_i \cdot W_j}{W} \cdot \frac{\overline{w}'}{n} = \Theta(\overline{w}'/n)\,. \end{aligned} \tag{4}$$

Note that this depends only on $i = \psi(u)$ and $j = \psi(v)$ and is independent of u and v themselves. This property does not hold for the Laplacian normalization above. It holds for the unnormalized adjacency matrix of the planted partition model based on $G_{n,p}$. In this case we have an expected value of $d_{ij} \cdot p = d_{ij} \cdot \overline{w}/n = \Theta(\overline{w}'/n)$. This is the first important analogy to the $G_{n,p}$–based model. The factor of $\overline{w}'^2$ in our normalization is only to see the analogies more clearly.

As for the adjacency matrix in the $G_{n,p}$-based model the spectrum of our matrix is soiled by rows, whose sum is considerably larger than their expectation. We remove all vertices from the graph (and of course the corresponding entries in M), whose row-sum in our normalized matrix M exceeds $C_1 \cdot \overline{w}'$ (Step 4 of Algorithm 1). After deleting these vertices the constructed matrix M^* allows to find the planted partition:

Theorem 4. *With high probability we have for all* $1 \leq i,j \leq 2$ *simultaneously:*

1. $\frac{\mathbf{1}^t}{\|\mathbf{1}\|} \cdot {M^*}_{V_i \times V_j} \cdot \frac{\mathbf{1}}{\|\mathbf{1}\|} = d_{ij} \cdot \frac{W_i \cdot W_j}{W} \cdot \sqrt{|V_i| \cdot |V_j|} \cdot \frac{\overline{w}'}{n} \cdot \left(1 \pm O\left(\frac{1}{\sqrt{\overline{w}'}}\right)\right).$
2. *For any* u, v *with* $\|u\| = \|v\| = 1$ *and* $u \perp \mathbf{1}$ *or* $v \perp \mathbf{1}$ *we have*

$$\left|u^t \cdot {M^*}_{V_i \times V_j} \cdot v\right| = O\left(\sqrt{C_1 \cdot \overline{w}'}\right).$$

For the intuition of 1., we refer to (4). Note,

$$\mathbf{E}\left[\frac{\mathbf{1}^t}{\|\mathbf{1}\|} \cdot M_{V_i \times V_j} \cdot \frac{\mathbf{1}}{\|\mathbf{1}\|}\right] = \frac{\mathbf{E}\left[s_M(V_i, V_j)\right]}{\sqrt{|V_i| \cdot |V_j|}} = d_{ij} \cdot \frac{W_i \cdot W_j}{W} \cdot \frac{\overline{w}'}{n} \cdot \sqrt{|V_i| \cdot |V_j|}$$

and so, the first item in Theorem 4 should be read as a concentration result.

Theorem 4 shows another analogy to the $G_{n,p}$-based model, see [3]. For unit-vectors u' and v' maximizing the term ${u'}^t \cdot {M^*}_{V_i \times V_j} \cdot v'$ we have that both u' and

v' are almost parallel to $\mathbf{1}$ and ${u'}^t M^* v' = \Theta(\overline{w}')$. Whereas if $u \perp u'$ or $v \perp v'$ then $u^t \cdot {M^*}_{V_i \times V_j} \cdot v$ is substantially smaller, namely $O(\sqrt{C_1 \cdot \overline{w}'})$.

The theorem above is the heart of our analysis. We will prove it in Section 4. In the next section we prove Theorem 3 by using Theorem 4.

3 Proof of Theorem 3

We start with a lemma about the eigenvalues of M^*. Its correctness is based mainly on Theorem 4 and the Courant-Fischer characterization of eigenvalues:

Fact 5. *Let $A \in \mathbb{R}^{n\times n}$ be some symmetric matrix with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$. Then for all $0 \leq j < n$*

$$\lambda_{j+1} = \min_{\substack{S \\ \dim S = j}} \max_{\substack{x \in S^\perp \\ \|x\|=1}} x^t A x$$

$$\lambda_{n-j} = \max_{\substack{S \\ \dim S = j}} \min_{\substack{x \in S^\perp \\ \|x\|=1}} x^t A x$$

where $S^\perp$ denotes the orthogonal complement to the subspace S.

Lemma 6. *With high probability M^* has exactly two eigenvalues, whose absolute value is $\Theta(\overline{w}')$, whereas all the other eigenvalues are $O\left(\sqrt{C_1 \cdot \overline{w}'}\right)$ in absolute value.*

Proof. Let U be the set constructed in Step 2. of our algorithm. Let χ_1 resp. χ_2 be $|U|$-dimensional characteristic vectors of $V_1 \cap U$ resp. $V_2 \cap U$ (the u-th component $\chi_i(u) = 1$ if $u \in V_i \cap U$ and 0 otherwise). We consider two vectors g and h from the space spanned by χ_1 and χ_2. Namely, $g = a_1 \cdot \chi_1/\|\chi_1\| + a_2 \cdot \chi_2/\|\chi_2\|$ with $a_1^2 + a_2^2 = 1$ and $h = b_1 \cdot \chi_1/\|\chi_1\| + b_2 \cdot \chi_2/\|\chi_2\|$ with $b_1^2 + b_2^2 = 1$. Note, $\|g\| = \|h\| = 1$. By Theorem 4 we have with probability $1 - o(1)$ that

$$\begin{aligned}
h^t M^* g &= \sum_{i,j=1}^{2} b_i \cdot \frac{\chi_i}{\|\chi_i\|} \cdot M^* \cdot a_j \cdot \frac{\chi_j}{\|\chi_j\|} = \sum_{i,j=1}^{2} b_i \cdot a_j \cdot \frac{\mathbf{1}^t \cdot {M^*}_{V_i \times V_j} \cdot \mathbf{1}}{\sqrt{|V_i \cap U| \cdot |V_j \cap U|}} \\
&= \sum_{i,j=1}^{2} b_i \cdot a_j \cdot d_{ij} \cdot \frac{W_i \cdot W_j}{W} \cdot \overline{w}' \cdot \frac{\sqrt{|V_i| \cdot |V_j|}}{n} \cdot \left(1 \pm O\left(\frac{1}{\sqrt{\overline{w}'}}\right)\right) \\
&= \sum_{i,j=1}^{2} \left(b_i \cdot a_j \cdot d_{ij} \cdot \frac{W_i \cdot W_j}{W} \cdot \overline{w}' \cdot \frac{\sqrt{|V_i| \cdot |V_j|}}{n}\right) \pm O\left(\sqrt{\overline{w}'}\right) \\
&= \frac{\overline{w}'}{W} \cdot \begin{pmatrix} b_1 \; b_2 \end{pmatrix} \cdot P \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \pm O\left(\sqrt{\overline{w}'}\right)
\end{aligned}$$

with

$$P = \begin{pmatrix} W_1 \cdot \sqrt{\frac{|V_1|}{n}} & 0 \\ 0 & W_2 \cdot \sqrt{\frac{|V_2|}{n}} \end{pmatrix} \cdot \begin{pmatrix} d_{11} \; d_{12} \\ d_{12} \; d_{22} \end{pmatrix} \cdot \begin{pmatrix} W_1 \cdot \sqrt{\frac{|V_1|}{n}} & 0 \\ 0 & W_2 \cdot \sqrt{\frac{|V_2|}{n}} \end{pmatrix}.$$

Remember, D has full rank as well as both remaining factors of P. We conclude that the matrix P has full rank. The W_i are $\Theta(1)$ as $|V_i|/n$, too. This shows that the spectral properties of P are determined only by D, ε and δ and do not rely on $w_1, \ldots, w_n$ or n. So P has two eigenvectors with constant nonzero eigenvalues. Let $\left(e_1\ e_2\right)^t$ and $\left(f_1\ f_2\right)^t$ be two orthonormal eigenvectors of P to the eigenvalues λ_1 and λ_2. Set

$$g_1 = e_1 \cdot \frac{\chi_1}{\|\chi_1\|} + e_2 \cdot \frac{\chi_2}{\|\chi_2\|} \qquad \text{and} \qquad g_2 = f_1 \cdot \frac{\chi_1}{\|\chi_1\|} + f_2 \cdot \frac{\chi_2}{\|\chi_2\|}.$$

By the calculation above get

$$\begin{aligned}\left|g_1^t \cdot M^* \cdot g_1\right| &= \left|\frac{\overline{w}'}{W} \cdot (e_1\ e_2) \cdot P \cdot \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \pm O\left(\sqrt{\overline{w}'}\right)\right| \\ &= \left|\frac{\overline{w}'}{W} \cdot \lambda_1 \pm O\left(\sqrt{\overline{w}'}\right)\right| = \Theta(\overline{w}')\end{aligned}$$

whereas

$$\begin{aligned}\left|g_1^t \cdot M^* \cdot g_2\right| &= \left|\frac{\overline{w}'}{W} \cdot (e_1\ e_2) \cdot P \cdot \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \pm O\left(\sqrt{\overline{w}'}\right)\right| \\ &= \left|\frac{\overline{w}'}{W} \cdot 0 \pm O\left(\sqrt{\overline{w}'}\right)\right| = O\left(\sqrt{\overline{w}'}\right).\end{aligned}$$

Thus for $1 \leq i, j \leq 2$ we have

$$\left|g_i^t \cdot M^* \cdot g_j\right| = \begin{cases} \Theta(\overline{w}') & \text{for } i = j \\ O\left(\sqrt{\overline{w}'}\right) & \text{for } i \neq j \end{cases}. \tag{5}$$

By Fact 5 we obtain, that at least two eigenvalues of M^* are $\Omega(\overline{w}')$ in absolute value: We subdivide g_1, g_2 by the sign of $g_i^t M^* g_i$. Let $g_1, \ldots, g_l$, $0 \leq l \leq 2$, be the vectors for which the product is positive, and $g_{l+1}, \ldots, g_2$ be those with $g_i^t M^* g_i < 0$. Let v be some unit-vector inside the subspace spanned by $g_1, \ldots, g_l$. We can rewrite $v = \sum_{i=1}^{l} \alpha_i \cdot g_i$ with $\sum_{i=1}^{l} \alpha_i^2 = 1$. Thus, by (5) there are two constants c and C such that

$$\begin{aligned} v^t \cdot M^* \cdot v &= \sum_{i,j=1}^{l} \alpha_i \alpha_j \cdot g_i^t M^* g_j = \sum_{i=1}^{l} \alpha_i^2 \cdot g_i^t M^* g_i + \sum_{\substack{i,j=1 \\ i \neq j}}^{l} \alpha_i \alpha_j \cdot g_i^t M^* g_j \\ &\geq \sum_{i=1}^{l} \alpha_i^2 \cdot c \cdot \overline{w}' \ - \ C \cdot \sqrt{\overline{w}'} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^{l} \alpha_i \alpha_j \\ &\geq c \cdot \overline{w}' \ - \ C \cdot \sqrt{\overline{w}'} \cdot l = \Omega(\overline{w}') \end{aligned}$$

The second equation of Fact 5 gives

$$\lambda_l = \max_{\substack{S \\ \dim S = |U| - l}} \ \min_{\substack{x \in S^\perp \\ \|x\| = 1}} x^t M^* x.$$

We choose S to be the orthogonal complement of $\langle g_1, \ldots, g_l \rangle$ which has dimension $|U| - l$. Together with the calculation above we get

$$\lambda_l \geq \min_{\substack{x \in \langle g_1, \ldots, g_l \rangle \\ \|x\|=1}} x^t M^* x \geq c \cdot \overline{w}' \; - \; C \cdot \sqrt{\overline{w}'} \cdot l = \Omega(\overline{w}').$$

To prove that $2-l$ eigenvalues are smaller than $-\Omega(\overline{w}')$, we use $g_{l+1}, \ldots, g_2$ and the first equation of Fact 5. So, we have at least 2 eigenvalues of M^* that are $\Omega(\overline{w}')$ in absolute value.

It is important that all the other eigenvalues of M^* are substantially smaller than $\overline{w}'$. Otherwise, it is hard to read the partition from the eigenvectors. Let u, v be any unit-vectors with u perpendicular to g_1 and g_2. Because both g_i are linear combinations of χ_1 and χ_2, u is also perpendicular to χ_1 and χ_2. Using Theorem 4 again, we obtain

$$\left| u^t M^* v \right| = \left| \sum_{i,j=1}^{2} u_{V_i \cap U} \cdot {M^*}_{V_i \times V_j} \cdot v_{V_j \cap U} \right| \leq 4 \cdot O\left(\sqrt{C_1 \cdot \overline{w}'} \right) \tag{6}$$

and in the same way $|v^t M^* u| = O\left(\sqrt{C_1 \cdot \overline{w}'} \right)$. The first equation of Fact 5 gives

$$\lambda_{l+1} \leq \max_{\substack{x \in \langle g_1, \ldots, g_l \rangle^{\perp} \\ \|x\|=1}} x^t M^* x$$

Let x be the vector maximizing the right-handed side. We rewrite the $x = \alpha \cdot u + \beta \cdot v$ with $u \perp g_1, g_2$ and $v \in \langle g_{l+1}, \ldots, g_2 \rangle$ and $\alpha^2 + \beta^2 = 1$. By the choice of l we have $v^t M^* v < 0$. With (6) we get

$$\lambda_{l+1} = x^t M^* x = \alpha^2 \cdot u^t M^* u + 2 \cdot \alpha\beta \cdot u^t M^* v + \beta^2 \cdot v^t M^* v = O(\sqrt{C_1 \cdot \overline{w}'})\,.$$

Using equation 2 of Fact 5 we obtain similarly

$$\lambda_{|U|-(2-l)} \geq \min_{\substack{x \in \langle g_{l+1}, \ldots, g_2 \rangle^{\perp} \\ \|x\|=1}} x^t M^* x \geq -C \cdot \sqrt{C_1 \cdot \overline{w}'}$$

for some constant $C > 0$. So, the remaining $|U| - 2$ eigenvalues of M^* are $O\left(\sqrt{C_1 \cdot \overline{w}'} \right)$ in absolute value. □

With Lemma 6 we can prove Theorem 3. Let e with $\|e\| = \sqrt{n}$ be an eigenvector of M^* with eigenvalue of size $\Theta(\overline{w}')$ (in absolute value). We can rewrite e as

$$e = \alpha \cdot \chi_1 + \beta \cdot \chi_2 + \gamma \cdot u \tag{7}$$

with $\|e\| = \|u\| = \sqrt{n}$. Again χ_1, χ_2 are the $|U|$-dimensional characteristic vectors of $V_1 \cap U$, $V_2 \cap U$ and $u \perp \chi_1, \chi_2$. As χ_1, χ_2 and u are pairwise orthogonal α, β and γ are unique. Note, that α and β can exceed 1. Theorem 4 yields

$|e^t \cdot M^* \cdot u| = O\left(n \cdot \sqrt{C_1 \cdot \overline{w}'}\right)$ since $u \perp \chi_1, \chi_2$ and $\|e\| = \|u\| = \sqrt{n}$. As $e^t \cdot M^* = \Theta(\overline{w}') \cdot e^t$ we get

$$O\left(n \cdot \sqrt{C_1 \cdot \overline{w}'}\right) = |e^t \cdot M^* \cdot u| = \Theta(\overline{w}') \cdot |e^t \cdot u| = \Theta(\overline{w}') \cdot |n \cdot \gamma|$$

leading to

$$|\gamma| = O\left(\sqrt{C_1/\overline{w}'}\right). \tag{8}$$

So, by the small value of γ u's impact on e is small.

In the remaining considerations of this section we will often use phrases like "*almost all* vertices in V_1 fulfill X". This means that the number of vertices in V_1 not satisfying X is bounded by $O(C_1 \cdot n/\overline{w}')$.

Let e be as above and $|\alpha - \beta| \geq 1/16$ then almost all vertices $v \in U$ have

$$|e(v) - \alpha| \leq \frac{1}{128} \quad \text{for } v \in V_1 \qquad \text{and} \qquad |e(v) - \beta| \leq \frac{1}{128} \quad \text{for } v \in V_2. \tag{9}$$

For any $v \in U$ that does not satisfy (9) we have $|\gamma \cdot u(v)| \geq 1/128$ and by (8) $|u(v)| = \Omega(\sqrt{\overline{w}'/C_1})$. Each of these entries contribute $\Omega(\overline{w}'/C_1)$ to $n = u^t \cdot u$. By this, the number of such entries is bounded above by $O(C_1 \cdot n/\overline{w}')$.

Let $s = \alpha \cdot \chi_1 + \beta \cdot \chi_2 + \gamma \cdot u$ be the vector determined by the algorithm. Then there are c_1, c_2 with $|c_1 - c_2| > 1/4$ and more than $n \cdot \sqrt{C_1/\overline{w}'}$ entries v in s fulfill

$$|s(v) - c_1| \leq \frac{1}{32} \qquad \text{resp.} \qquad |s(v) - c_2| \leq \frac{1}{32}. \tag{10}$$

We distinguish two cases. We start with $|\alpha - \beta| \geq 1/16$. As more than $n \cdot \sqrt{C_1/\overline{w}'}$ vertices fulfill (10) and almost all vertices fulfill (9) at least one $v \in V_1$ agrees both (9) and (10), provided $\overline{w}'$ is large enough. Assume, that v fulfills the first inequality in (10). By the triangle inequality we get

$$|\alpha - c_1| \leq |\alpha - s(v)| + |s(v) - c_1| \leq \frac{1}{128} + \frac{1}{32} = \frac{5}{128}$$

and by the same argument $|\beta - c_2| \leq 5/128$. Clearly, it is also possible that $|\alpha - c_2| \leq 5/128$ and $|\beta - c_1| \leq 5/128$. However, this situation can be handled analogously. Because of (9) we have that almost all $v \in V_1$ have

$$|s(v) - c_1| \leq |s(v) - \alpha| + |\alpha - c_1| \leq \frac{1}{128} + \frac{5}{128} = \frac{3}{64}$$

and almost all entries $v \in V_2$ have $|s(v) - c_2| \leq 3/64$. Since $|c_1 - c_2| \geq 16/64$ Algorithm 2 classifies almost all vertices correctly.

Now we come to the case $|\alpha - \beta| < 1/16$. Assume c_1 is farther away from $(\alpha + \beta)/2$ as c_2 is. As $|c_1 - c_2| > 1/4$ the distance of c_1 to $(\alpha + \beta)/2$ is at least $1/8$. Since $|\alpha - \beta| \leq 1/16$ we have

$$|\alpha - c_1| \geq \frac{1}{8} - \frac{1}{32} = \frac{3}{32} \qquad \text{and} \qquad |\beta - c_1| \geq \frac{3}{32}.$$

As more than $n \cdot \sqrt{C_1/\overline{w}'}$ components v of s have $|s(v) - c_1| \le 1/32$ the same components in u fulfill $|\gamma \cdot u(v)| \ge 2/32$. By (8) we have that $u(v) \ge \Omega(\sqrt{\overline{w}'/C_1})$, so

$$n = u^t \cdot u \ge n \cdot \sqrt{\frac{C_1}{\overline{w}'}} \cdot \Omega\left(\frac{\overline{w}'}{C_1}\right) = \Omega\left(n \cdot \sqrt{\overline{w}'/C_1}\right),$$

which is a contradiction for large enough $\overline{w}'$. So, the vector s chosen in the algorithm has $|\alpha - \beta| \ge 1/16$.

We have shown above that the vector s yields a good approximation of the planted partition, provided Lemma 6 holds. We are left to show that a vector s agreeing the requirements stated in Step 6. exists w.h.p. Let s_1, s_2 as in the algorithm and $s_i = \alpha_i \cdot \chi_1 + \beta_i \cdot \chi_2 + \gamma_i \cdot u_i$ its decomposition as in (7).

Assume for a contradiction $|\alpha_i - \beta_i| \le 1/4$ for both $i = 1, 2$. As

$$n = s_i^t \cdot s_i = \alpha_i^2 \cdot |V_1 \cap U| + \beta_i^2 \cdot |V_2 \cap U| + \gamma_i^2 \cdot n$$

we obtain by dividing through n

$$\alpha_i^2 + \beta_i^2 \ge \alpha_i^2 \cdot \frac{|V_1 \cap U|}{n} + \beta_i^2 \cdot \frac{|V_2 \cap U|}{n} = 1 - \gamma_i^2 \overset{(8)}{\ge} 1 - O\left(C_1/\overline{w}'\right). \qquad (11)$$

Now it is clear that $|\alpha_i| > 1/2$ or $|\beta_i| > 1/2$ holds. Using the assumption $|\alpha_i - \beta_i| \le 1/4$ we see that α_i and β_i have the same sign. So

$$|\alpha_1 \cdot \alpha_2 + \beta_1 \cdot \beta_2| = |\alpha_1 \cdot \alpha_2| + |\beta_1 \cdot \beta_2| \ge \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{2} = 1/4$$

leading to

$$\begin{aligned} 0 &= s_1^t \cdot s_2 = \left|\alpha_1 \cdot \alpha_2 \cdot |V_1 \cap U| + \beta_1 \cdot \beta_2 \cdot |V_2 \cap U| + \gamma_1 \cdot \gamma_2 \cdot u_1^t \cdot u_2\right| \\ &\ge (\delta n - |V \setminus U|) \cdot |\alpha_1 \cdot \alpha_2 + \beta_1 \cdot \beta_2| - |\gamma_1 \cdot \gamma_2| \cdot n \\ &\overset{(8),(17)}{\ge} n \cdot (\delta \cdot |\alpha_1 \cdot \alpha_2 + \beta_1 \cdot \beta_2| - O(C_1/\overline{w}')) \ge n \cdot (\delta/4 - O\left(C_1/\overline{w}'\right)). \end{aligned}$$

As δ is some positive constant and $\overline{w}'$ is large, we obtain a contradiction. So, at least one s_i has $|\alpha_i - \beta_i| > 1/4$. Inequality (9) shows that this s_i complies the requirements of Step 6 with $c_1 := \alpha_i$ and $c_2 := \beta_i$. As s_i is $|U|$-dimensional at least

$$|U| - O(C_1 \cdot n/\overline{w}') \overset{(17)}{\ge} n - O(C_1 \cdot n/\overline{w}')$$

vertices are classified correctly. □

4 Proof of Theorem 4: The spectrum of $M^*_{V_i \times V_j}$

As stated in (3) each entry of the submatrix $M_{V_i \times V_j}$ has the same expected value. Besides we know that each entry has two possible values: 0 and some positive real number. However the non-zero values of the entries typically differ from each other, because we multiply with $\overline{w}'^2/(w'_u \cdot w'_v)$. Nonetheless we know that the maximum entry in M is bounded above by $\overline{w}'^2/(\min_{u \in V} w'_u)^2$. We summarize all these facts in

Definition 7. *We call a real $n \times m$-matrix $X = (x_{uv})$ a* same-mean-*matrix with mean μ and bound b if the following conditions hold*

1. *The x_{uv} are independent random variables (the trivial dependence induced by symmetry is allowed).*
2. *Each x_{uv} has exactly two possible values, one of both is 0.*
3. *There is a bound b such that definitely $x_{uv} \le b$ for all u, v.*
4. *$\mathbf{E}[x_{uv}] = \mu > 0$ for all u, v.*

Note, item 3. needs $x_{uv} \le b$ independently of the concrete outcome X.

It is not hard to check that $M_{V_i \times V_j}$ is a same-mean-matrix with mean

$$\mu \overset{(4)}{=} d_{ij} \cdot \frac{W_i \cdot W_j}{W} \cdot \frac{\overline{w}'}{n} \qquad \text{and} \quad \text{bound} \;\; b = \frac{\overline{w}'^2}{(\min_u w'_u)^2} = \Theta(1). \tag{12}$$

Unfortunately, ${M^*}_{V_i \times V_j}$ does not have property 1 stated in Definition 7 as we deleted some vertices. So, we concentrate on $M_{V_i \times V_j}$ and transfer the necessary results to ${M^*}_{V_i \times V_j}$.

The following lemma is important to the analysis of same-mean-matrices. It is a generalization of Lemma 3.4 in [3] and can be proven in a similar way as Alon and Kahale did it.

Lemma 8. *Let X be some same-mean-matrix with mean μ and bound b and $y_1, \ldots, y_l$ be a set of mutually independent entries of it. Let $a_1, \ldots, a_l$ be arbitrary real numbers from the interval $[-a, a]$. If S, D and some constant $c > 0$ fulfill*

$$\sum_{i=1}^{l} a_i^2 \le D \qquad \text{and} \qquad S \le c \cdot \mathrm{e}^c \cdot D \cdot \mu / a,$$

then the new random variable $Z = \sum_{i=1}^{l} a_i \cdot y_i$ fulfills

$$\mathbf{Pr}\left[|Z - \mathbf{E}[Z]| \ge S\right] \le 2 \cdot \exp\bigl(-S^2/(2\mu \cdot \mathrm{e}^c \cdot D \cdot b)\bigr)\,.$$

Proof. In case of $b = 1$, the lemma can be proven in the same way as Lemma 3.4 in [3] for random 0-1-variables. The details can be found in Subsection 5.1.

We come to the case $b \ne 1$. Let $X = (x_{uv})$ be the matrix in Lemma 8. We construct the matrix $X' = (x'_{uv})$ by setting $x'_{uv} := x_{uv}/b$. Clearly, all x'_{uv} are independent and bounded above by 1. Their expectation $\mathbf{E}[x'_{uv}]$ is $\mathbf{E}[x_{uv}]/b = \mu/b$. So, X' is a same-mean-matrix with mean μ/b and bound 1.

Let $y_1, \ldots, y_l$ be as in the assertion and $y'_1, \ldots, y'_l$ be the corresponding entries in X'. Then $Z' = \sum_{i=1}^{l} a_i \cdot y'_i = Z/b$ and as $S \le c \cdot \mathrm{e}^c \cdot D \cdot \mu/a$,

$$S' = \frac{S}{b} \le \frac{c \cdot \mathrm{e}^c \cdot D \cdot \mu/a}{b} = c \cdot \mathrm{e}^c \cdot D \cdot \mu'/a\,.$$

We finish the proof with an application of the case $b = 1$ on X'

$$\begin{aligned}\mathbf{Pr}\left[|Z - \mathbf{E}[Z]| \geq S\right] = \mathbf{Pr}\left[\left|\frac{Z}{b} - \frac{\mathbf{E}[Z]}{b}\right| \geq \frac{S}{b}\right] &= \mathbf{Pr}\left[|Z' - \mathbf{E}[Z']| \geq S'\right] \\ &\leq 2 \cdot \exp\big(-S'^2/(2\mu' \cdot \mathrm{e}^c \cdot D)\big) \\ &= 2 \cdot \exp\Big(-(S/b)^2 \ / \ (2 \cdot (\mu/b) \cdot \mathrm{e}^c \cdot D)\Big) \\ &= 2 \cdot \exp\big(-S^2/(2 \cdot \mu \cdot \mathrm{e}^c \cdot D \cdot b)\big)\,.\end{aligned}$$

□

4.1 The first item of Theorem 4

To determine $\mathbf{1}^t \cdot {M^*}_{V_i \times V_j} \cdot \mathbf{1} = s_M(V_i \cap U, V_j \cap U)$ it suffices to subtract the sum of the entries we delete in Step 4 of our algorithm from $s_M(V_i, V_j)$, the sum of all entries in $M_{V_i \times V_j}$.

The latter one can be determined by Lemma 8, as $M_{V_i \times V_j}$ is a same-mean-matrix with $\mu = d_{ij} \cdot W_i \cdot W_j / W \cdot \overline{w}'/n$ and bound $\Theta(1)$, see (12). In case $i \neq j$ all entries of $M_{V_i \times V_j}$ are independent. If we choose all the a's in Lemma 8 to 1, $D = |V_i| \cdot |V_j|$ and $c = \ln 4$, we see

$$\begin{aligned}\mathbf{Pr}\left[|s_M(V_i, V_j) - \mu \cdot |V_i| \cdot |V_j|| \geq \mu \cdot |V_i| \cdot |V_j| / \sqrt{\overline{w}'}\right] &\leq 2\exp(-\mu \cdot D/(8 \cdot \overline{w}' \cdot b)) \\ &= \exp(-\Omega(n))\end{aligned}$$

A similar equation can be obtained for $i = j$. The trivial symmetry in the entries does no harm. We simply use only the upper triangle of the matrix to get: With high probability

$$s_M(V_i, V_j) = \mu \cdot |V_i| \cdot |V_j| \cdot (1 \pm O(1/\sqrt{\overline{w}'})). \tag{13}$$

Now we bound the sum of the entries we delete in Step 4. As a first step we bound the *number* of vertices we delete. The row-sum $s_M(u, V)$ of such an entry $u \in V_i$ has to be larger than $C_1 \cdot \overline{w}'$ with $C_1 \geq 5/W \cdot \max_{i,j}\{d_{ij} \cdot W_i \cdot W_j\} = \Theta(1)$ by (2). In contrast the expected row-sum is

$$\mathbf{E}[s_M(u, V)] = W_i \cdot \frac{\overline{w}'}{W \cdot n} \cdot \left(\sum_{v \in V_1} d_{i1} \cdot W_1 + \sum_{v \in V_2} d_{i2} \cdot W_2\right) \leq \frac{C_1}{5} \cdot \overline{w}'\,. \tag{14}$$

So we have a deviation from the expectation by a factor of at least 4 which is very unlikely as Lemma 8 shows. Note, we have to bound $s_M(u, V_1)$ and $s_M(u, V_2)$ seperately as the expectation of the entries differs. If $s_M(u, V) \geq 5 \cdot \mathbf{E}[s_M(u, V)]$ then $s_M(u, V_1) \geq 5 \cdot \mathbf{E}[s_M(u, V_1)]$ or $s_M(u, V_2) \geq 5 \cdot \mathbf{E}[s_M(u, V_2)]$. Let the a_i's in Lemma 8 be 1, $D = |V_i|$ and $c = \ln 4$. We obtain for fixed u

$$\mathbf{Pr}\left[\ |s_M(u, V_i) - \mathbf{E}[s_M(u, V_i)]|\ \geq 4 \cdot \mathbf{E}[s_M(u, V_i)]\ \right] \leq 2 \cdot \exp(-c_{\mathrm{m}} \cdot \overline{w}'/b)\ , \tag{15}$$

where

$$c_{\mathrm{m}} = \delta/W \cdot \min_{\substack{i,j\\ d_{ij}>0}} (d_{ij} \cdot W_i \cdot W_j) = \Theta(1)\,. \tag{16}$$

Note that $c_{\mathrm{m}} \cdot \overline{w}'$ represents a lower bound on the minimal expected row-sum in any (non-zero) $M_{V_i \times V_j}$. As $b = \Theta(1)$ by (12) we replace c_{m}/b with $c = \Theta(1)$.

By (15), the expected number of vertices not belonging to U is bounded by $4 \cdot \exp(-c \cdot \overline{w}') \cdot n$. We can use Chebycheff's inequality to show that with probability $1 - o(1)$

$$|V \setminus U| \leq 8 \cdot \exp(-c \cdot \overline{w}') \cdot n\,. \tag{17}$$

Now we bound the *sum* of the entries we delete from $M_{V_i \times V_j}$, namely

$$s_M(V_i \setminus U, V_j \setminus U) + s_M(V_i \cap U, V_j \setminus U) + s_M(V_i \setminus U, V_j \cap U).$$

The entries of M are non-negative so any upper bound on $s_M(V_i \setminus U, V_j) + s_M(V_i, V_j \setminus U)$ is an upper bound on the sum above.

We show the first summand in detail. As we do not know U's size exactly, we consider all sets $X \subset V_i$ with $|X| = 8 \cdot \exp(-c \cdot \overline{w}') \cdot n$. Clearly, $V_i \setminus U$ is a subset of at least one such X, because $|V_i \setminus U| \leq |V \setminus U| \leq 8 \cdot \exp(-c \cdot \overline{w}') \cdot n$ and it suffices to show an upper bound on $s_M(X, V_j)$. We have

$$\mathbf{E}\left[s_M(X, V_j)\right] = d_{ij} \cdot \frac{W_i \cdot W_j}{W} \cdot \frac{\overline{w}'}{n} \cdot |X| \cdot |V_j| = \mu \cdot |X| \cdot |V_j|\,.$$

In case $i = j$ we have some small dependencies because of symmetry of the entries. This concerns only few entries, because of the relatively small size of X. Using only the independent entries for Lemma 8 we get

$$\begin{aligned}\mathbf{Pr}\left[\, s_M(X, V_j) \geq 10 \cdot \mathbf{E}\left[s_M(X, V_j)\right] \,\right] &\leq 2 \cdot \exp(-2 \cdot \mu \cdot |X| \cdot |V_j|)\\ &\leq 2 \cdot \exp(-2 \cdot \mu \cdot |X| \cdot \delta n)\\ &\leq 2 \cdot \exp(-2 \cdot c_{\mathrm{m}} \cdot \overline{w}' \cdot |X|)\ .\end{aligned}$$

We have only

$$\binom{|V_i|}{|X|} \leq \binom{n}{|X|} \leq \left(\frac{\mathrm{e} \cdot n}{|X|}\right)^{|X|} \leq \exp(c \cdot \overline{w}' \cdot |X|)$$

different sets X of size $8 \cdot \exp(-c \cdot \overline{w}') \cdot n$. As $b \geq 1$, $c = c_{\mathrm{m}}/b \leq c_{\mathrm{m}}$ a simple union bound yields that with probability at least

$$1 - 2 \cdot \exp(-c_{\mathrm{m}} \cdot \overline{w}' \cdot |X|) = 1 - 2 \cdot \exp\Big(-8 \cdot c_{\mathrm{m}} \cdot \overline{w}' \cdot \mathrm{e}^{-c \cdot \overline{w}'} \cdot n\Big) = 1 - o(1)$$

all sets X as above fulfill $s_M(X, V_j) \leq 80 \cdot \exp(-c \cdot \overline{w}') \cdot n$. As the same bound holds w.h.p. for $s_M(V_i, Y)$ for all $Y \subset V_j$ with $|Y| = 8 \cdot \exp(-c \cdot \overline{w}') \cdot n$, the sum of the entries inside $M_{V_i \times V_j}$ we deleted is bounded above by

$$s_M(V_i \setminus U, V_j) + s_M(V_i, V_j \setminus U) \leq 160 \cdot \mathrm{e}^{-c \cdot \overline{w}'} \cdot n \tag{18}$$

with probability $1-o(1)$. We conclude by (13) and (18), the term $\mathbf{1}^t \cdot {M^*}_{V_i \times V_j} \cdot \mathbf{1} = s_M(V_i \cap U, V_j \cap U)$ is bounded below by

$$\mu \cdot |V_i| \cdot |V_j| \cdot (1 \pm O(1/\sqrt{\overline{w}'})) - 160 \mathrm{e}^{-c \cdot \overline{w}'} \cdot n = \mu \cdot |V_i| \cdot |V_j| \cdot (1 \pm O(1/\sqrt{\overline{w}'}))$$

for $\overline{w}'$ large enough as $\mu = \Theta(\overline{w}'/n)$ and $|V_i|, |V_j| \geq \delta n = \Omega(n)$. Finally we get

$$\begin{aligned} \frac{\mathbf{1}^t}{\|\mathbf{1}\|} \cdot {M^*}_{V_i \times V_j} \cdot \frac{\mathbf{1}}{\|\mathbf{1}\|} &= \frac{\mathbf{1}^t \cdot {M^*}_{V_i \times V_j} \cdot \mathbf{1}}{\sqrt{|V_i \cap U| \cdot |V_j \cap U|}} \\ &= \frac{\mu \cdot |V_i| \cdot |V_j| \cdot (1 \pm O(1/\sqrt{\overline{w}'}))}{\sqrt{|V_i| \cdot |V_j|} \cdot (1 - 8 \cdot \exp(-c \cdot \overline{w}'))^2} \\ &= \mu \cdot \sqrt{|V_i| \cdot |V_j|} \cdot (1 \pm O(1/\sqrt{\overline{w}'})) \,. \end{aligned}$$

The claim follows immediately as $\mu = d_{ij} \cdot W_i \cdot W_j / W \, \cdot \, \overline{w}'/n$. □

4.2 The second item of Theorem 4

Using the techniques of [14] and [3] together with Lemma 8 we can prove

Lemma 9. *Let X be an $n \times m$–same-mean-matrix with mean μ and bound b and $N = n + m$. Let $R = \{u : \sum_v x_{uv} \leq d \cdot \mu \cdot N\}$ and $C = \{v : \sum_u x_{uv} \leq d \cdot \mu \cdot N\}$ for $d > 1$ arbitrary.*

If $\mu \cdot n \cdot m > b \cdot N$, then we have with probability $1 - O(1/\sqrt{N})$ for all pairs of vectors u and v, with $\|u_{|R}\| = \|v_{|C}\| = 1$ and $u_{|R} \perp \mathbf{1}$ or $v_{|C} \perp \mathbf{1}$

$$\left| {u_{|R}}^t \cdot X \cdot v_{|C} \right| = O(\sqrt{b \cdot d \cdot \mu \cdot N}).$$

Proof. In conjunction with Lemma 8 the proof for the case $b = 1$ is strongly related to the proof of Lemma 3.3 in [3] respectively Theorem 2.2 in [14]. We postpone the proof for $b = 1$ to Subsection 5.2.

We are left to show the case $b \neq 1$. We rewrite X as $X = b \cdot X'$. Then, X' is a same-mean-matrix with mean μ/b and bound 1. Note, the sets R and C are the same for X and X' and all conditions are fulfilled. We apply Lemma 9 for $b = 1$ to X'.

$$\begin{aligned} \left| {u_{|R}}^t \cdot X \cdot v_{|C} \right| &= \left| {u_{|R}}^t \cdot b \cdot X' \cdot v_{|C} \right| = b \cdot \left| {u_{|R}}^t \cdot X' \cdot v_{|C} \right| = b \cdot O(\sqrt{d \cdot \mu/b \cdot N}) \\ &= O(\sqrt{b \cdot d \cdot \mu \cdot N}). \end{aligned}$$

□

Let u be some $|V_i|$-dimensional vector and v be some $|V_j|$-dimensional vector. Clearly we have

$$u_U^t \cdot {M^*}_{V_i \times V_j} \cdot v_U = {u^t}_{|U} \cdot M_{V_i \times V_j} \cdot v_{|U},$$

as u_U is the vector where the entries $\notin U$ are deleted and $u_{|U}$ is the vector where these entries are set to 0.

We want to bound ${u^t}_{|U} \cdot M_{V_i \times V_j} \cdot v_{|U}$ using Lemma 9. For an application we have to check that $M_{V_i \times V_j}$ agrees the conditions of Lemma 9 and we need $U \cap V_i \subseteq R$ and $U \cap V_j \subseteq C$.

Remember, $M_{V_i \times V_j}$ is a same-mean matrix with bound $\overline{w}'^2/(\min_u w'_u)^2 = O(1)$, see (12). U contains only those vertices whose row-sum (and by symmetry whose column-sum, too) in M is at most $C_1 \cdot \overline{w}'$. Let $d := C_1/c_{\mathrm{m}}$, see (16) for c_{m}'s value. Then for each $u \in U \cap V_i$

$$d \cdot \mu \cdot N = \frac{C_1}{c_{\mathrm{m}}} \cdot \mu \cdot N \geq \frac{C_1}{c_{\mathrm{m}}} \cdot d_{ij} \cdot \frac{W_i \cdot W_j}{W} \cdot \frac{\overline{w}'}{n} \cdot (2\delta n) \geq 2 \cdot C_1 \cdot \overline{w}' > \sum_{v \in V_j} x_{uv}$$

and we see $U \cap V_i \subseteq R$ and analogously $U \cap V_j \subseteq C$. Since $\mu \cdot |V_i| \cdot |V_j| = \Omega(\overline{w}' \cdot n)$ and $b \cdot N = b \cdot (|V_i| + |V_j|) = O(n)$ all conditions of Lemma 9 are fulfilled.

Proving item 2 of Theorem 4 is now easy. Any vector pair u, v as in the theorem can be extended to some $|V_i|$-dimensional vector u' (resp. $|V_j|$-dimensional vector v') by filling up with 0's. Clearly, $\|u\| = \|u'\| = 1$ and $\|v\| = \|v'\| = 1$. If $u \perp \mathbf{1}$ (here $\mathbf{1}$ is $|V_i \cap U|$-dimensional) then ${u'}_{|R} \perp \mathbf{1}$. We can use Lemma 9 to bound

$$\begin{aligned} u^t \cdot {M^*}_{V_i \times V_j} \cdot v &= {u'}_{|U} \cdot M_{V_i \times V_j} \cdot {v'}_{|U} = {u'}_{|R} \cdot M_{V_i \times V_j} \cdot {v'}_{|C} \\ &= O\left(\sqrt{b \cdot C_1/c_{\mathrm{m}} \cdot \mu \cdot (|V_i| + |V_j|)}\right) = O\left(\sqrt{C_1 \cdot \overline{w}'}\right). \end{aligned}$$

□

5 Technical lemmas

5.1 Proof of Lemma 8 for $b = 1$

The proof follows the proof of Lemma 3.4 in [3]. We omitted the bounds of any $\prod$ and any $\sum$ as in the whole section the index i passes through $1, \ldots, l$.

Let $p_1, \ldots, p_l$ be the probabilities of $y_1, \ldots, y_l$ being non-zero. Thus $p_i \cdot y_i = \mu$ if y_i is non-zero. In other words, the second value y_i can have (besides 0) is μ/p_i. With Markov's inequality we get

$$\begin{aligned} \mathbf{Pr}\left[Z - \mathbf{E}\left[Z\right] \geq S\right] &= \mathbf{Pr}\left[\mathrm{e}^{\lambda(Z - \mathbf{E}[Z])} \geq \mathrm{e}^{\lambda S}\right] = \mathbf{Pr}\left[\mathrm{e}^{\lambda(Z - \mathbf{E}[Z] - S)} \geq 1\right] \\ &\leq \mathbf{E}\left[\mathrm{e}^{\lambda(Z - \mathbf{E}[Z] - S)}\right] = \frac{\mathbf{E}\left[\exp(\lambda Z)\right]}{\exp(\lambda(\mathbf{E}\left[Z\right] + S))}. \end{aligned} \tag{19}$$

By setting $\lambda = S/(\mathrm{e}^c \cdot \mu \cdot D) \leq c/a$ we get $\lambda \cdot a_i \cdot y_i \leq \lambda \cdot a_i \leq c$. Remember, no y_i exceeds 1 as X is a same-mean-matrix with bound 1. We take a closer look at the enumerator of (19):

$$
\begin{aligned}
\mathbf{E}\left[\exp(\lambda Z)\right] &= \mathbf{E}\left[\exp\Big(\lambda \sum a_i \cdot y_i\Big)\right] = \mathbf{E}\left[\prod \exp(\lambda a_i \cdot y_i)\right] \\
&= \prod \mathbf{E}\left[\exp(\lambda a_i \cdot y_i)\right] \\
&= \prod \Big(p_i \cdot \exp(\lambda \cdot a_i \cdot \mu/p_i) + (1-p_i) \cdot \exp(\lambda \cdot a_i \cdot 0) \Big) \\
&= \prod \big(1 + p_i \cdot \big(\exp(\lambda \cdot a_i \cdot \mu/p_i) - 1\big)\big) \,.
\end{aligned}
$$

Now we use that the function $f(x) = e^c \cdot x^2/2 + x - (e^x - 1)$ is convex for $x \le c$ and the minimum in the interval $(-\infty, c]$ is $f(0) = 0$. So f is non-negative for all $x \le c$ and $e^x - 1 \le x + e^c \cdot x^2/2$ for all $x \le c$. For $x = \lambda \cdot a_i \cdot \mu/p_i$ we have $x \le c$ as $y_i \le \mu/p_i \le 1$. Then we get

$$
\begin{aligned}
\mathbf{E}\left[\exp(\lambda Z)\right] &\le \prod \left(1 + p_i \left(\lambda \cdot a_i \cdot \mu/p_i + \frac{\mathrm{e}^c}{2} \cdot (\lambda \cdot a_i \cdot \mu/p_i)^2\right)\right) \\
&\le \prod \left(1 + \lambda \cdot a_i \cdot \mu + \mathrm{e}^c \cdot \lambda^2 \cdot a_i^2 \cdot \mu^2/(2 \cdot p_i)\right) \\
&\le \prod \left(1 + \lambda \cdot a_i \cdot \mu + \mathrm{e}^c \cdot \lambda^2 \cdot a_i^2 \cdot \mu/2\right) \,.
\end{aligned}
$$

Since $1 + x \le e^x$ for all $x \in \mathbb{R}$ we have

$$
\begin{aligned}
\mathbf{E}\left[\exp(\lambda Z)\right] &\le \prod \exp\big(\lambda \cdot a_i \cdot \mu + \mathrm{e}^c \cdot \lambda^2 \cdot a_i^2 \cdot \mu/2\big) \\
&= \exp\Big(\sum \lambda \cdot a_i \cdot \mu + \mathrm{e}^c \cdot \lambda^2 \cdot a_i^2 \cdot \mu/2\Big) \\
&= \exp\Big(\lambda \cdot \mathbf{E}\left[Z\right] + \mathrm{e}^c \cdot \lambda^2 \cdot \mu/2 \cdot \sum a_i^2\Big) \\
&\le \exp\big(\lambda \cdot \mathbf{E}\left[Z\right] + \mathrm{e}^c \cdot \lambda^2 \cdot \mu \cdot D/2\big) \\
&= \exp(\lambda \cdot \mathbf{E}\left[Z\right] + \lambda \cdot S/2) \,.
\end{aligned}
$$

Summing up, we get for (19)

$$
\begin{aligned}
\mathbf{Pr}\left[Z - \mathbf{E}\left[Z\right] \ge S\right] &\le \frac{\mathbf{E}\left[\exp(\lambda Z)\right]}{\exp(\lambda(\mathbf{E}\left[Z\right] + S))} \le \frac{\exp(\lambda \cdot \mathbf{E}\left[Z\right] + \lambda \cdot S/2)}{\exp(\lambda(\mathbf{E}\left[Z\right] + S))} \\
&\le \exp(-\lambda \cdot S/2) = \exp\left(\frac{-S^2}{2 \cdot e^c \cdot \mu \cdot D}\right) .
\end{aligned}
$$

By negating all a_i's, we can obtain $\mathbf{Pr}\left[\mathbf{E}\left[Z\right] - Z \ge S\right] \le \exp\Big(\frac{-S^2}{2 \cdot \mathrm{e}^c \cdot \mu \cdot D}\Big)$ in the same way. □

5.2 Proof of Lemma 9 for $b = 1$

The proof of Lemma 9 follows the ideas of [14] and [3]. In the whole section all O- and Ω-terms are based on N and hold for all $N > N_0 =$ constant.

Lemma 10. *Let* $X = (x_{uv})$ *be some same-mean-matrix with mean* μ *and bound* 1. *Then with probability* $1 - O(1/\sqrt{N})$ *for any pair* (A, B) *of sets* $A \subseteq [n]$, $B \subseteq [m]$ *the following holds: If* $K = \max\{|A|, |B|\} \leq N/2$ *then*

1. $s_X(A, B) \leq 200 \cdot \mathbf{E}[s_X(A, B)]$ or
2. $s_X(A, B) \cdot \ln \frac{s_X(A,B)}{\mathbf{E}[s_X(A,B)]} \leq 200 \cdot K \cdot \ln \frac{N}{K}$.

is fulfilled.

Proof. We will prove the lemma for symmetric matrices because due to the dependence of the entries by symmetry it is slightly harder to show.

Fix two sets A and B and set $\eta = \mathbf{E}[s_X(A, B)] = |A| \cdot |B| \cdot \mu$. There is an unique number β such that

$$\beta \cdot \ln \beta = 200 \cdot K \cdot \ln(N/K)/\eta.$$

Then condition 2 equals $s_X(A, B) \leq \beta \cdot \eta$. We set $\beta' = \max\{200, \beta\}$ and it suffices to show that, with high probability, no pair (A, B) with $s_X(A, B) > \beta' \cdot \eta$ exists. For that we want to use Lemma 8. Due to symmetry not all random variables are independent. For $u \neq v$ both in $A \cap B$ the entries x_{uv} and x_{vu} are equal and obviously dependent. In that case we use only x_{uv} with $u < v$ and assign the corresponding a_i in Lemma 8 to 2, because x_{uv} is counted twice in $s_X(A, B)$.

For the other pairs $(u, v) \in A \times B$ we assign a_i to 1, whereby a can be 2. The value of $D = \sum a_i^2$ lies between $|A| \cdot |B|$ and $2 \cdot |A| \cdot |B|$. Choose c such that $c \cdot \mathrm{e}^c = \beta' - 1$. We get for fixed A and B

$$\begin{aligned}
\mathbf{Pr}[s_X(A, B) \geq \beta' \cdot \eta] &= \mathbf{Pr}[s_X(A, B) \geq (c \cdot \mathrm{e}^c + 1) \cdot \eta] \\
&\leq \mathbf{Pr}[s_X(A, B) - \eta \geq c \cdot \mathrm{e}^c \cdot \eta] \\
&\leq \mathbf{Pr}[|s_X(A, B) - \eta| \geq c \cdot \mathrm{e}^c \cdot |A| \cdot |B| \cdot \mu] \\
&\leq \mathbf{Pr}[|s_X(A, B) - \eta| \geq c \cdot \mathrm{e}^c \cdot D \cdot \mu/a] \\
&\leq 2 \cdot \exp\left(-\frac{(c \cdot \mathrm{e}^c \cdot D \cdot \mu/a)^2}{2 \cdot \mu \cdot \mathrm{e}^c \cdot D}\right) \leq 2 \cdot \exp\left(-\frac{c^2 \cdot \mathrm{e}^c \cdot D \cdot \mu}{2 \cdot a^2}\right) \\
&\leq 2 \cdot \exp\left(-\frac{c^2 \cdot \mathrm{e}^c \cdot \eta}{8}\right) = 2 \cdot \exp(-c \cdot (\beta' - 1) \cdot \eta/8)\ .
\end{aligned}$$

Since $\beta' = c \cdot e^c + 1 \geq 200$ we have $c > 3$. By this we can bound c from below by $\ln \beta'/2$. Then we get further

$$\begin{aligned}
\mathbf{Pr}[s_X(A, B) \geq \beta' \cdot \eta] &\leq 2\exp(-\ln \beta' \cdot (\beta' - 1) \cdot \eta/16) \\
&\leq 2\exp(-\ln \beta' \cdot \beta' \cdot \eta/32) \\
&\leq 2\exp(-200 \cdot K \cdot \ln(N/K)/32) = 2\left(\frac{K}{N}\right)^{200 \cdot K/32} \\
&\leq 2\left(\frac{K}{N}\right)^{6K}
\end{aligned}$$

since $K < N$. The number of pairs (A, B) possible is bounded above by

$$\sum_{i=1}^{K} 2 \cdot \binom{N}{K} \cdot \binom{N}{i} \leq 2K \cdot \binom{N}{K}^2 \leq 2K \cdot \left(\frac{\mathrm{e} \cdot N}{K}\right)^{2K}.$$

So, the probability that any pair (A, B) with $\max\{|A|, |B|\}$ fulfills $s_X(A, B) \geq \beta' \cdot \eta$ is bounded above by

$$2\left(\frac{K}{N}\right)^{6K} \cdot 2K \cdot \left(\frac{\mathrm{e} \cdot N}{K}\right)^{2K} = 4K \cdot \left(\mathrm{e}^2 \cdot \frac{K^4}{N^4}\right)^{K} < 4K \cdot \left(\frac{\mathrm{e}^2}{8} \cdot \frac{K}{N}\right)^{K} = O\left(\frac{K}{N}\right)^{K}.$$

Summing over all possible values for K we get a bound of

$$\begin{aligned}
\sum_{K=1}^{N/2} \left(\frac{K}{N}\right)^K &\leq \sum_{K=1}^{\sqrt{N}} \left(\frac{K}{N}\right)^K + \sum_{K=\sqrt{N}}^{N/2} \left(\frac{K}{N}\right)^K \\
&\leq \sum_{K=1}^{\sqrt{N}} \left(\frac{1}{\sqrt{N}}\right)^K + \sum_{K=\sqrt{N}}^{\infty} \left(\frac{1}{2}\right)^K \\
&\leq \frac{1}{\sqrt{N}} + \sqrt{N} \cdot \frac{1}{N} + 2 \cdot \left(\frac{1}{2}\right)^{\sqrt{N}} \leq \frac{2}{\sqrt{N}} + 2 \cdot \left(\frac{1}{2}\right)^{\sqrt{N}} \\
&= O(1/\sqrt{N})
\end{aligned}$$

□

Now we come to the proof of Lemma 9. We assume that Lemma 10 holds, as it does w.h.p. The proof of the case $u_{|R} \perp \mathbf{1}$ is given in detail. The case $v_{|C} \perp \mathbf{1}$ can be treated identically.

Note, there is an uncountable number of vectors u and v over $\mathbb{R}$. To tackle this problem we approximate the vectors over $\mathbb{R}$ we consider by the following ε-nets.

$$\begin{aligned}
T_n &= \left\{ x \in \left(\frac{\mathbb{Z}}{2\sqrt{n}}\right)^n : \|x\| \leq 2 \right\} \\
T_R &= \left\{ x_{|R} : x \in T_n \right\} \\
T_C &= \left\{ x_{|C} : x \in T_m \right\}.
\end{aligned}$$

Note, $T_R \subseteq T_n$ and $T_C \subseteq T_m$. One can show that $|T_n| \leq k^n$ for some constant k and: For u, v as in the assumption there are $l \in T_R$ and $r \in T_C$ with $l \perp \mathbf{1}$ such that

$$\left| {u_{|R}}^t \cdot X \cdot v_{|C} \right| \leq 5 \cdot \left| l^t \cdot X \cdot r \right|.$$

For details, see [14] or [9, Section 5.4.5], for example.

So, in order to prove Lemma 9, it suffices to show that for all l and r as in the inequality above $|l^t \cdot X \cdot r|$ is bounded by $O(\sqrt{d \cdot \mu \cdot N})$. For the remaining part of this section we only deal with l and r instead of u and v. All occurrences of u (resp. v) refers to indices between 1 and n (resp. m). Let $l \in T_R$, $r \in T_C$ with $l \perp \mathbf{1}$. In order to bound

$$\left|l^t \cdot X \cdot r\right| = \left|\sum_{u,v=1}^{u=n,v=m} l_u x_{uv} r_v\right|$$

we define

$$\mathcal{B} = \mathcal{B}(l,r) = \left\{(u,v) : |l_u r_v| \le \sqrt{d \cdot \mu / N}\right\}.$$

We will show that with probability $1 - O(1/\sqrt{N})$

$$\left|\sum_{(u,v)\in\mathcal{B}} l_u x_{uv} r_v\right| = O(\sqrt{d \cdot \mu \cdot N}) \qquad \text{and} \qquad \left|\sum_{(u,v)\notin\mathcal{B}} l_u x_{uv} r_v\right| = O(\sqrt{d \cdot \mu \cdot N}).$$

We start with the "small" pairs, namely those in $\mathcal{B}$. Fix two vectors $l \in T^n$ and $r \in T^m$ with $l \perp \mathbf{1}$. Clearly, $\sum_{u=1}^n l_u = 0$. So

$$\sum_{(u,v)} l_u \cdot r_v = \sum_{v=1}^{m} r_v \sum_{u=1}^{n} l_u = 0 \quad \text{and} \quad \sum_{(u,v)\in B} l_u r_v = -\sum_{(u,v)\notin B} l_u r_v,$$

and we can bound $\left|\sum_{(u,v)\in B} l_u r_v\right|$ by $\left|\sum_{(u,v)\notin B} l_u r_v\right|$. We have

$$\begin{aligned}
\left|\sum_{(u,v)\notin\mathcal{B}} l_u r_v\right| &= \left|\sqrt{\frac{N}{d\cdot\mu}} \cdot \sum_{(u,v)\notin\mathcal{B}} \sqrt{d\cdot\mu/N} \cdot l_u r_v\right| \\
&= \sqrt{\frac{N}{d\cdot\mu}} \cdot \left|\sum_{(u,v)\notin\mathcal{B}} \sqrt{d\cdot\mu/N} \cdot l_u r_v\right| \le \sqrt{\frac{N}{d\cdot\mu}} \cdot \left|\sum_{(u,v)\notin\mathcal{B}} l_u^2 r_v^2\right| \\
&\le \sqrt{\frac{N}{d\cdot\mu}} \cdot \sum_{u=1}^{n} l_u^2 \cdot \sum_{v=1}^{m} r_v^2 \le 16 \cdot \sqrt{N/(d\cdot\mu)}\,.
\end{aligned}$$

as $|l_u r_v| \ge \sqrt{d \cdot \mu / N}$ for $(u,v) \notin \mathcal{B}$. Bounding the expectation of $\sum\limits_{(u,v)\in\mathcal{B}} l_u x_{uv} r_v$ is now easy:

$$\left|\mathbf{E}\left[\sum_{(u,v)\in\mathcal{B}} l_u x_{uv} r_v\right]\right| \le \left|\sum_{(u,v)\in\mathcal{B}} l_u \cdot \mathbf{E}\left[x_{uv}\right] r_v\right| = \left|\mu \cdot \sum_{(u,v)\in\mathcal{B}} l_u r_v\right| \le 16 \cdot \sqrt{\frac{\mu \cdot N}{d}}.$$

Next we want to bound the probability of a large deviation of $\sum_{(u,v)\in\mathcal{B}} l_u x_{uv} r_v$ from its expectation. Obviously, the term is a weighted sum of entries of the same-mean-matrix X and Lemma 8 may help.

Yet, Lemma 8 requires the chosen entries y_i to be independent. This is not the case if both (u,v) and (v,u) belong to $\mathcal{B}$. In that case we can join $l_u x_{uv} r_v$ and $l_v x_{vu} r_u$ to $x_{uv}(l_u r_v + l_v r_u)$, because $x_{uv} = x_{vu}$.

So, the y_i's for the lemma are the x_{uv}'s with $(u,v) \in \mathcal{B}$ (resp. x_{uv} with $u < v$, if $(u,v) \in \mathcal{B}$ and $(v,u) \in \mathcal{B}$), the a_i's are the $l_u \cdot r_v$ (resp. $l_u r_v + l_v r_u$). The bound a for all a_i is $2 \cdot \sqrt{\mu/N}$ by the definition of $\mathcal{B}$. To set $D = 64 \cdot d > 64$ ensures $\sum a_i^2 \leq 4 \cdot \sum l_u^2 \cdot r_v^2 \leq 64 \leq D$. We get that

$$\mathbf{Pr}\left[\left|\sum_{(u,v)\in\mathcal{B}} l_u x_{uv} r_v - \mathbf{E}\left[\sum_{(u,v)\in\mathcal{B}} l_u x_{uv} r_v\right]\right| \geq 32 \cdot c \cdot \mathrm{e}^c \cdot \sqrt{d \cdot \mu \cdot N}\right] \leq 2\exp\bigl(-8 \cdot c^2 \cdot \mathrm{e}^c \cdot N\bigr)$$

for any constant $c > 0$. There are at most

$$|T_R| \cdot |T_C| \leq |T_n| \cdot |T_m| \leq k^{n+m} = k^N = \exp(N \cdot \ln k)$$

possible vector-pairs (l,r). A simple union bound yields for sufficiently large (but still constant) c that with probability $1 - O(1/N)$ for *all* $l \in T_R$, $r \in T_C$, with $l \perp \mathbf{1}$, simultaneously we have

$$\left|\sum_{(u,v)\in\mathcal{B}} l_u x_{uv} r_v\right| = O(\sqrt{d \cdot \mu \cdot N})\,.$$

We are left to show the bound for the "large" pairs being not in $\mathcal{B}$. For this we subdivide *all* entries of l and r (without restricting to $\mathcal{B}$). Namely let for $i > 0$

$$A_i = \left\{u : \frac{2^{i-1}}{2\sqrt{N}} \leq l_u < \frac{2^i}{2\sqrt{N}}\right\} \quad \text{and} \quad A_i = \left\{u : \frac{2^{|i|-1}}{2\sqrt{N}} \leq -l_u < \frac{2^{|i|}}{2\sqrt{N}}\right\}$$

for $i < 0$. Let $a_i = |A_i|$ for all i.

Note, we have $O(\log N)$ non-empty sets A_i, otherwise $\|l\|$ would exceed 2. There is no need to define A_0. Each entry l_u smaller than $2^0/(2\sqrt{N})$ and larger than $-2^0/(2\sqrt{N})$ must be 0 by the definition of T_n as $N > n$. Such entries have no impact to the following calculations. Define B_j and b_j analogously for r.

We use the notation $i \sim j$ if $2^{|i|+|j|}/4 > \sqrt{d \cdot \mu \cdot N}$. For $u \in A_i$, $v \in B_j$ and $(u,v) \notin \mathcal{B}$ we have

$$\frac{2^{|i|}}{2\sqrt{N}} > |l_u|\,, \qquad \frac{2^{|j|}}{2\sqrt{N}} > |r_v| \qquad \text{and} \qquad |l_u r_v| > \sqrt{d \cdot \mu/N}.$$

So $2^{|i|+|j|}/(4N)$ must be larger than $\sqrt{d \cdot \mu/N}$, so $2^{|i|+|j|}/4 > \sqrt{d \cdot \mu \cdot N}$ yielding $i \sim j$. By this we can bound

$$\left|\sum_{(u,v)\notin\mathcal{B}} l_u x_{uv} r_v\right| \leq \sum_{(u,v)\notin\mathcal{B}} |l_u x_{uv} r_v| \leq \sum_{i\sim j} \sum_{\substack{u\in A_i\\ v\in B_j}} |l_u x_{uv} r_v|\,.$$

We can split the last term into eight sums separating by the signs of i and j and the fact whether $a_i \geq b_j$ or $a_i < b_j$. Let

$$\mathcal{C} = \{(i,j) : i \sim j,\ i,j > 0, a_i < b_j\}.$$

As the proofs for all eight sums are very similar, we give only the proof for

$$\sum_{(i,j)\in\mathcal{C}} \sum_{\substack{u\in A_i\\ v\in B_j}} |l_u x_{uv} r_v| = O(\sqrt{d \cdot \mu \cdot N}).$$

To make the following calculations clearer, we need some abbreviations:

$$s_{ij} = s_X(A_i, B_j) \qquad \lambda_{ij} = \frac{s_{ij}}{\mu_{ij}} \qquad \alpha_i = \frac{a_i \cdot (2^i)^2}{4N}$$

$$\mu_{ij} = \mathbf{E}\left[s_{ij}\right] = a_i \cdot b_j \cdot \mu \qquad \sigma_{ij} = \frac{\lambda_{ij} \cdot \sqrt{d \cdot \mu \cdot N}}{2^{i+j-2}} \qquad \beta_j = \frac{b_j \cdot (2^j)^2}{4N}$$

Note, that λ_{ij} denotes the relative deviation of $s_X(A_i, B_j)$ from its expectation μ_{ij}. σ_{ij} is merely a technical term. Since $i \sim j$ we have $\sigma_{ij} < \lambda_{ij}$ and σ_{ij}/λ_{ij} becomes small if we deal with very large pairs ($|l_u r_v| \gg \sqrt{d \cdot \mu/N}$). The term α_i bounds $\sum_{u\in A_i} l_u^2$ as

$$\alpha_i/4 \leq \sum_{u\in A_i} l_u^2 < \alpha_i.$$

Summing over all i yields $\sum_i \alpha_i \leq 4 \cdot \|l\|^2 \leq 16$. Clearly, in the same way we get $\sum_j \beta_j \leq 16$, too. For $i,j > 0$ we can bound

$$\begin{aligned}
\sum_{\substack{u\in A_i\\ v\in B_j}} |l_u x_{uv} r_v| &\leq \sum_{\substack{u\in A_i\\ v\in B_j}} \frac{2^{i+j}}{4N} \cdot x_{uv} = \frac{2^{i+j}}{4N} \cdot s_{ij} = \frac{2^{i+j}}{4N} \cdot \lambda_{ij} \cdot \mu_{ij} \\
&= \frac{2^{i+j}}{4N} \cdot \frac{\sigma_{ij} \cdot 2^{i+j-2}}{\sqrt{d \cdot \mu \cdot N}} \cdot \mu_{ij} = \frac{2^{2i} \cdot 2^{2j}}{16N} \cdot \frac{\sigma_{ij}}{\sqrt{d \cdot \mu \cdot N}} \cdot a_i \cdot b_j \cdot \mu \\
&= \alpha_i \cdot \beta_j \cdot \sigma_{ij} \cdot \sqrt{\mu \cdot N/d}\,.
\end{aligned}$$

So it suffices to show $\sum_{(i,j)\in\mathcal{C}} \alpha_i \cdot \beta_j \cdot \sigma_{ij} = O(d)$.

If we transfer Lemma 10 to the notation given above, we obtain that with high probability for each A_i, B_j

$$\lambda_{ij} \leq 200 \tag{20}$$

or

$$\sigma_{ij} \cdot \alpha_i \cdot \log \lambda_{ij} \leq \frac{200 \cdot 2^{i-j} \cdot \sqrt{d}}{\sqrt{\mu \cdot N}} \cdot (2j - \log \beta_j) \tag{21}$$

hold.

Note, pairs (A_i, B_j) with $b_j > N/2$ are not covered by the lemma. Yet in that case we have $\lambda_{ij} < 2d$. This can be seen as follows: $l_u = 0$ if $s_X(u,V) > d \cdot \mu \cdot N$ and u occurs in none of the A_i's. So, $s_X(u,V) \leq d \cdot \mu \cdot N$ for all $u \in A_i$. This leads to

$$s_{ij} \leq d \cdot \mu \cdot N \cdot a_i. \tag{22}$$

If $b_j > N/2$ then $\mu_{ij} > a_i \cdot N/2 \cdot \mu$, yielding $\lambda_{ij} = s_{ij}/\mu_{ij} < 2d$.

We subdivide the pairs $(i,j) \in \mathcal{C}$ into six classes $\mathcal{C}_1, \ldots, \mathcal{C}_6$ so that $(i,j) \in \mathcal{C}_k$ if (i,j) fulfills the following condition k, but none of the conditions $< k$.

1. $\lambda_{ij} \leq 200d$
2. $\sigma_{ij} \leq 1$
3. $2^{i-j} > \sqrt{d \cdot \mu \cdot N}$
4. $\log \lambda_{ij} \geq (2j - \log \beta_j)/4$ and $2j > -\log \beta_j$
5. $\log \lambda_{ij} < (2j - \log \beta_j)/4$ and $2j > -\log \beta_j$
6. $2j \leq -\log \beta_j$

If we can prove for each $\mathcal{C}_k$ that $\sum_{(i,j)\in\mathcal{C}_k} \alpha_i \beta_j \sigma_{ij} = O(d)$, we are done.

1. $\lambda_{ij} \leq 200d$

   Since $\sigma_{ij} < \lambda_{ij}$ we get easily

$$\begin{aligned}\sum_{(i,j)\in\mathcal{C}_1} \alpha_i\beta_j\sigma_{ij} &= \sum_{(i,j)\in\mathcal{C}_1} 200d \cdot \alpha_i\beta_j \leq 200d \cdot \sum_{(i,j)} \alpha_i\beta_j \\ &= 200d \cdot \sum_i \alpha_i \cdot \sum_j \beta_j \leq 200d \cdot 16 \cdot 16 = O(d).\end{aligned}$$

2. $\sigma_{ij} \leq 1$

   Analogously to 1. we obtain $\sum_{(i,j)\in\mathcal{C}_2} \alpha_i\beta_j\sigma_{ij} \leq 16 \cdot 16 = O(1)$.

3. $2^{i-j} > \sqrt{d \cdot \mu \cdot N}$

   We have $s_{ij} \leq d \cdot \mu \cdot N \cdot a_i$ by (22) and

$$s_{ij} = \lambda_{ij} \cdot \mu_{ij} = \sigma_{ij} \cdot 2^{i+j-2} \cdot a_i \cdot b_i \cdot \sqrt{\mu/(d \cdot N)}.$$

   Both together give

$$\begin{aligned}\sigma_{ij} \cdot 2^{i+j-2} \cdot b_i \cdot \sqrt{\mu}/\sqrt{d \cdot N} &\leq d \cdot \mu \cdot N \\ \sigma_{ij} \cdot 2^{i+j-2} \cdot b_i/N &\leq \sqrt{d^3 \cdot \mu \cdot N} \\ \sigma_{ij} \cdot 2^{2j} \cdot b_i/(4N) &\leq \sqrt{d^3 \cdot \mu \cdot N} \cdot 2^{j-i} \\ \sigma_{ij}\beta_j &\leq \sqrt{d^3 \cdot \mu \cdot N} \cdot 2^{j-i}.\end{aligned}$$

So we have

$$\begin{aligned}
\sum_{(i,j)\in\mathcal{C}_3} \alpha_i\beta_j\sigma_{ij} &\le \sum_{(i,j)\in\mathcal{C}_3} \alpha_i \cdot \sqrt{d^3\cdot\mu\cdot N}\cdot 2^{j-i} \\
&\le \sum_{\substack{(i,j)\in\mathcal{C}\\ j<i-\log\sqrt{d\cdot\mu\cdot N}}} \alpha_i \cdot \sqrt{d^3\cdot\mu\cdot N}\cdot 2^{j-i} \\
&= \sum_{i>\log\sqrt{d\cdot\mu\cdot N}} \alpha_i \sum_{j=1}^{i-\log\sqrt{d\cdot\mu\cdot N}-1} \sqrt{d^3\cdot\mu\cdot N}\cdot 2^{j-i} \\
&\le \sum_{i>\log\sqrt{d\cdot\mu\cdot N}} \alpha_i \cdot \sqrt{d^3\cdot\mu\cdot N}\cdot (2^{-\log\sqrt{d\cdot\mu\cdot N}} - 1/2^i) \\
&< \sum_{i>\log\sqrt{d\cdot\mu\cdot N}} \alpha_i \cdot d = O(d)\,.
\end{aligned}$$

For the three remaining cases we use inequality (21). In these cases (as also in 2. and 3.) $\lambda_{ij} > 200d > 200$, so inequality (20) is violated and (21) must hold.

4. $\log\lambda_{ij} \ge (2j - \log\beta_j)/4$ and $2j > -\log\beta_j$
From inequality (21) and $\log\lambda_{ij} \ge (2j-\log\beta_j)/4$ we obtain

$$\sigma_{ij}\cdot\alpha_i\cdot(2j-\log\beta_j)/4 \le \frac{200\cdot 2^{i-j}\cdot\sqrt{d}}{\sqrt{\mu\cdot N}}\cdot(2j-\log\beta_j)$$

and as $2j - \log\beta_j$ is positive

$$\sigma_{ij}\cdot\alpha_i \le \frac{800\cdot 2^{i-j}\cdot\sqrt{d}}{\sqrt{\mu\cdot N}}.$$

We know $2^{i-j} \le \sqrt{d\cdot\mu\cdot N}$ because we left behind $\mathcal{C}_3$. So $i \le j + \log\sqrt{d\cdot\mu\cdot N}$ and

$$\begin{aligned}
\sum_{(i,j)\in\mathcal{C}_4} \alpha_i\beta_j\sigma_{ij} &\le \sum_{j>0} \sum_{\substack{(i,j)\in\mathcal{C}\\ i\le j+\log\sqrt{d\cdot\mu\cdot N}}} \alpha_i\beta_j\sigma_{ij} \\
&\le \sum_{j>0}\beta_j \cdot \sum_{i=1}^{j+\log\sqrt{d\cdot\mu\cdot N}} \frac{800\cdot 2^{i-j}\cdot\sqrt{d}}{\sqrt{\mu\cdot N}} \\
&< \sum_{j>0}\beta_j\cdot\frac{800\cdot 2^{\log\sqrt{d\cdot\mu\cdot N}+1}\cdot\sqrt{d}}{\sqrt{\mu\cdot N}} = \sum_{j>0}\beta_j\cdot 1600d = O(d).
\end{aligned}$$

5. $\log\lambda_{ij} < (2j-\log\beta_j)/4$ and $2j > -\log\beta_j$
From (21) together with the latter condition we conclude

$$\sigma_{ij}\cdot\alpha_i \le \frac{200\cdot 2^{i-j}\cdot\sqrt{d}}{\sqrt{\mu\cdot N}}\cdot 4j. \tag{23}$$

Note, we omitted $\log \lambda_{ij} > \log(200) > 1$ on the left-handed side.
We have $\log \lambda_{ij} < (2j - \log \beta_j)/4 \leq j$. By the definition of σ_{ij} we obtain $\sigma_{ij} \leq \sqrt{d \cdot \mu \cdot N}/2^{i-2}$. The case $\sigma_{ij} \leq 1$ is handled above, so $2^{i-2} < \sqrt{d \cdot \mu \cdot N}$ holds. Now, we bound

$$\begin{aligned}
\sum_{(i,j)\in\mathcal{C}_5} \alpha_i \beta_j \sigma_{ij} &\leq \sum_{\substack{(i,j)\in\mathcal{C}_5 \\ i<\log\sqrt{d\cdot\mu\cdot N}+2}} \alpha_i \beta_j \sigma_{ij} \leq \sum_{j>0} \beta_j \sum_{i=1}^{\log\sqrt{d\cdot\mu\cdot N}+1} \alpha_i \sigma_{ij} \\
&\overset{(23)}{\leq} \sum_{j>0} \beta_j \sum_{i=1}^{\log\sqrt{d\cdot\mu\cdot N}+1} 800 \cdot j \cdot 2^{i-j} \cdot \sqrt{d}/\sqrt{\mu \cdot N} \\
&\leq \sum_{j>0} \beta_j \cdot 800 \cdot j \cdot 2^{-j} \cdot \sum_{i=1}^{\log\sqrt{d\cdot\mu\cdot N}+1} 2^i \cdot \sqrt{d}/\sqrt{\mu \cdot N} \\
&< \sum_{j>0} \beta_j \cdot 800 \cdot j \cdot 2^{-j} \cdot 4d \leq O(d) \cdot \sum_{j>0} \frac{j}{2^j} = O(d).
\end{aligned}$$

6. $2j \leq -\log \beta_j$
As $j > 0$ we have that $\beta_j < 1$ and $-\log \beta_j$ is positive. Again, we use (21) and omit $\log \lambda_{ij} > 1$:

$$\sigma_{ij} \cdot \alpha_i \leq -\frac{400 \cdot 2^{i-j} \cdot \sqrt{d}}{\sqrt{\mu \cdot N}} \cdot \log \beta_j.$$

Remember, since $2^{i-j} \leq \sqrt{d \cdot \mu \cdot N}$ (case 3. is handled above) we get $i \leq j + \log \sqrt{d \cdot \mu \cdot N}$ and so

$$\begin{aligned}
\sum_{(i,j)\in\mathcal{C}_6} \alpha_i \beta_j \sigma_{ij} &\leq \sum_{(i,j)\in\mathcal{C}_6} -\frac{400 \cdot 2^{i-j} \cdot \sqrt{d}}{\sqrt{\mu \cdot N}} \cdot \beta_j \cdot \log \beta_j \\
&\leq \sum_{j>0} -\beta_j \cdot \log \beta_j \cdot \sum_{i=1}^{j+\log\sqrt{d\cdot\mu\cdot N}} \frac{400 \cdot 2^{i-j} \cdot \sqrt{d}}{\sqrt{\mu \cdot N}} \\
&< \sum_{j>0} -\beta_j \cdot \log \beta_j \cdot 800d
\end{aligned}$$

As β_j can be very small, we need to bound $-\log \beta_j$. As $0 < \beta_j < 1$ we have $-\log \beta_j < 4/\sqrt{\beta_j}$. In order to prove this, we show that the function $f(x) = 4/\sqrt{x} + \log x$ is positive for $0 < x < 1$. The first derivative $f'(x) = (-2/\sqrt{x} + 1/\ln 2)/x$ is negative for $0 < x \leq 1$. So, $f(x)$ falls strictly monotonic in our interval. As $f(1) = 4 > 0$, our bound hold for all $0 < x \leq 1$.
So, $-\beta_j \cdot \log \beta_j < \beta_j/\sqrt{\beta_j} = \sqrt{\beta_j}$. At a next step we replace $\sqrt{\beta_j}$ by 2^{-j} since $2j < -\log \beta_j$ gives $2^{-2j} > \beta_j$. So

$$\sum_{(i,j)\in\mathcal{C}_6} \alpha_i \beta_j \sigma_{ij} \leq \sum_{j>0} 800d \cdot \sqrt{\beta_j} < 800d \cdot \sum_{j>0} 2^{-j} = 800d = O(d).$$

□

References

1. Aiello, W, Chung, F., Lu, L.: A random graph model for massive graphs. Proc. 33rd. SToC (2001), 171–180.
2. Alon, N. Spectral techniques in graph algorithms. Proc. LATIN (1998), LNCS 1380, Springer, 206–215.
3. Alon, N., Kahale, N.: A spectral technique for coloring random 3-colorable graphs. SIAM J. Comput. **26** (1997) 1733–1748.
4. Boppana, R.B.: Eigenvalues and graph bisection: An average case analysis. Proc. 28th FoCS (1987), 280–285.
5. Chung, F.K.R.: Spectral Graph Theory. American Mathematical Society (1997).
6. Chung, F.K.R., Lu, L., Vu, V.: The Spectra of Random Graphs with Given Expected Degrees. Internet Mathematics **1** (2003) 257–275.
7. Coja-Oghlan, A.: On the Laplacian eigenvalues of $G_{n,p}$. Preprint (2005) http://www.informatik.hu-berlin.de/∼coja/de/publikation.php.
8. Coja-Oghlan, A., Lanka, A.: The Spectral Gap of Random Graphs with Given Expected Degrees. Preprint (2006).
9. Coja-Oghlan, A.: Spectral techniques, semidefinite programs, and random graphs. Habilitationsschrift, Humboldt Universität zu Berlin, Institut für Informatik.
10. Coja-Oghlan, A.: An adaptive spectral heuristic for partitioning random graphs. Proc. ICALP (2006), 691–702.
11. Dasgupta, A., Hopcroft, J.E., McSherry, F.: Spectral Analysis of Random Graphs with Skewed Degree Distributions. Proc. 45th FOCS (2004) 602–610.
12. Feige, U., Ofek, E.: Spectral Techniques Applied to Sparse Random Graphs. Random Structures and Algorithms, **27(2)** (2005), 251–275.
13. Flaxman, A.: A spectral technique for random satisfiable 3CNF formulas. Proc. 14th SODA (2003) 357–363.
14. Friedman, J., Kahn, J., Szemerédi, E.: On the Second Eigenvalue in Random Regular Graphs. Proc. 21th STOC (1989) 587–598.
15. Füredi, Z., Komloś, J.: The eigenvalues of random symmetric matrices. Combinatorica **1** (1981) 233–241.
16. Giesen, J., Mitsche, D.: Reconstructing Many Partitions Using Spectral Techniques. Proc. 15th FCT (2005) 433–444.
17. Husbands, P., Simon, H., and Ding, C.: On the use of the singular value decomposition for text retrieval. In 1st SIAM Computational Information Retrieval Workshop (2000), Raleigh, NC.
18. Janson, S., Łuczak, T., Ruciński, A.: Random graphs. John Wiley and Sons 2000.
19. Krivelevich, M., Sudakov, B.: The largest eigenvalue of sparse random graphs. Combinatorics, Probability and Computing **12** (2003) 61–72.
20. Krivelevich, M., Vu, V.H.: On the concentration of eigenvalues of random symmetric matrices. Microsoft Technical Report 60 (2000).
21. Lempel, R., Moran, S. Rank-stability and rank-similarity of link-based web ranking algorithms in authority-connected graphs. Information retrieval, special issue on Advances in Mathematics/Formal methods in Information Retrieval (2004) Kluwer.
22. McSherry, F.: Spectral Partitioning of Random Graphs. Proc. 42nd FoCS (2001) 529–537.
23. Meila, M., Varna D.: A comparison of spectral clustering algorithms. UW CSE Technical report 03-05-01.

24. Mihail, M., Papadimitriou, C.H.: On the Eigenvalue Power Law. Proc. 6th RANDOM (2002) 254–262.
25. Pothen, A., Simon, H.D., Liou, K.-P.: Partitioning sparse matrices with eigenvectors of graphs. SIAM J. Matrix Anal. Appl. **11** (1990) 430–452
26. Spielman, D.A., Teng, S.-H.: Spectral partitioning works: planar graphs and finite element meshes. Proc. 36th FOCS (1996) 96–105.