# The spectrum of random $k$-lifts of large graphs (with possibly large $k$)[*]

Roberto Imbuzeiro Oliveira[†]

We study random $k$-lifts of large, but otherwise arbitrary graphs $G$. We prove that, with high probability, all eigenvalues of the adjacency matrix of the lift that are not eigenvalues of $G$ are of the order of $\sqrt{\Delta_G \ln(kn)}$, where $\Delta_G$ is the maximum degree of $G$. Similarly, and also with high probability, the "new" eigenvalues of the normalized Laplacian of $G^{(k)}$ are all in an interval of length $O(\sqrt{\ln(nk)/\delta_G})$ around 1, where $\delta_G$ is the minimum degree of $G$.

We also prove that, from the point of view of spectra, there is very little difference between a random $k_1 k_2 \ldots k_r$-lift of a graph and a random $k_1$-lift of a random $k_2$-lift of $\ldots$ of a random $k_r$-lift of the same graph.

The main proof tool is a concentration inequality for sums of independent random matrices which is of independent interest.

AMS 2000 subject classifications: Primary 05C80; secondary 05C80.
Keywords and phrases: Random lifts of graphs, graph spectrum.

## 1. Introduction

Let $G$ be a graph with vertex set $V$ and edge set $E$. A $k$-lift of $G$ is a graph $G^{(k)}$ with vertex set $V \times [k]$ and edge set:

$$E^{(k)} \equiv \cup_{vw \in E} \mathcal{M}_{vw}$$

where each $\mathcal{M}_{vw}$ is a matching of the sets $\{v\} \times [k]$ and $\{w\} \times [k]$. In more intuitive terms: each vertex of $G$ is replaced by $k$ copies of itself and each edge $vw \in E$ is replaced by a matching of the copies of $v$ and $w$.

There have been many recent results about *random $k$-lifts of graphs* where $G$ is fixed and $k \to +\infty$. Here "random" means that the matchings $\mathcal{M}_{vw}$ are chosen independently and each of them is uniformly distributed.

---

A lot is now known about properties of $G^{(k)}$ such as connectivity [2, 3], chromatic number [4], spectral distribution [10, 17] and the existence of perfect matchings [16].

A disjoint line of work has considered 2-lifts of arbitrary (possibly large) graphs $G$. The goal in this case was to provide an explicit construction of *some* 2-lift with good spectral properties, so that arbitrarily large expanders can be efficiently constructed via successive 2-lifts [5].

In this paper we study a scenario that is quite natural but mostly new: random $k$-lifts of large graphs $G$. We obtain non-trivial results only when the minimum degree of $G$ is $\gg \ln(|V|k)$, but $G$ and $k$ are otherwise arbitrary. For concrete examples, one may think of random $n$-lifts of graphs on $n$ vertices and minimal degree $\ln^{1+\epsilon} n$; or of $2^{\sqrt{n}}$-lifts of $(n/2)$-regular graphs on $n$ vertices. The only other work on random lifts of large graphs that we are aware of is the preprint by Lubetsky et al. [18], which covers regular $n$-vertex graphs with $n \ll k^{2/3}$ (more on this paper below); other cases (including constant $k$) seem to be unexploited.

Our focus will be on the spectra of the *adjacency matrix* and the normalized *Laplacian*, or simply *the Laplacian* (as opposed to the combinatorial Laplacian, cf. Remark 2.1), of the random lift. In this paper we will treat the spectra of these matrices as objects of intrinsic interest, but one should note that they can be used to estimate many parameters of graphs, including the diameter, distances between distinct subsets, discrepancy-like properties, path congestion, cuts, chromatic number and the mixing time for random walk; see e.g. [13, 14, 8].

We need some preliminaries in order to state our theorem. Let $A$ and $A^{(k)}$ be the adjacency matrix of the graph $G$ and of its $k$-lift $G^{(k)}$ (resp.). We will see in Section 3.1 that the spectrum of $A^{(k)}$ always contains the spectrum of $A$ in the sense of *multisets*: any eigenvalue of $A$ with multiplicity $m$ is an eigenvalue of $A^{(k)}$ with multiplicity $\geq m$. The same holds for the spectra of the normalized Laplacians $\mathcal{L}^{(k)}$ and $\mathcal{L}$ of $G^{(k)}$ and $G$ (respectively).

Let new$(A^{(k)})$ be the difference between the spectrum of $A^{(k)}$ and the spectrum of $A$ and define new$(\mathcal{L}^{(k)})$ similarly. Note that new$(A^{(k)})$ is also a multiset: if $\lambda$ has multiplicity $m_1$ in the spectrum of $A$ and multiplicity $m_2$ in the spectrum of $A^{(k)}$, it occurs $m_2 - m_1$ times in new$(A^{(k)})$.

Our main result concerns the new eigenvalues of *weakly random $k$-lifts*, a generalization of the standard random $k$-lifts discussed above; see Definition 3.1 for details.

**Theorem 1.1.** *Let $G = (V, E)$ be a graph on $n = |V|$ vertices with minimum (resp. maximum) degree $\delta_G$ (resp. $\Delta_G$). Let $A$ and $\mathcal{L}$ denote the adjacency*

*matrix and normalized Laplacian of $G$ and let $A^{(k)}$ and $\mathcal{L}^{(k)}$ denote the corresponding matrices for a weakly random $k$-lift of $G$ (cf. Definition 3.1). Then for all $\eta \in (0,1)$,*

$$\mathbb{P}\left(\sup_{\eta \in \mathrm{new}(A^{(k)})} |\eta| \leq 2\sqrt{2\Delta_G \ln(2nk/\eta)}\right) \geq 1 - \eta$$

*and*

$$\mathbb{P}\left(\sup_{\beta \in \mathrm{new}(\mathcal{L}^{(k)})} |1 - \beta| \leq 2\sqrt{\frac{2\ln(2nk/\eta)}{\delta_G}}\right) \geq 1 - \eta.$$

This is interesting even in the case $k = 2$. It is known [5] that any $d$-regular graph has a two-lift whose new eigenvalues are all bounded by $O(\sqrt{d \ln d})$. However, a *typical* random 2-lift of $G_n$ might have at least one new eigenvalue equal to $d$. One example (also from [5]) consists of $n/(d+1)$ disjoint $(d+1)$-cliques; the new eigenvalue $d$ comes from there being a clique whose lift consists of two disconnected cliques.[1] Notice that the probability of there being such a clique is $1 - o(1)$ even when $d = \lceil c\sqrt{\ln n} \rceil$ for some small constant $c > 0$. On the other hand, Theorem 1.1 shows that there exists some $C > 0$ such that for any $\epsilon > 0$, if $d \geq C \ln n/\epsilon^2$, then the largest new eigenvalue is $\leq \epsilon d$ with probability $\geq 1 - 1/n^2$.

Before we proceed, we note some limitations of our theorem. The largest eigenvalue of $A^{(k)}$ is always between $\delta_G$ and $\Delta_G$ and the eigenvalues of $\mathcal{L}^{(k)}$ are always between 0 and 2 [13]. Hence our result for the adjacency matrix is trivial if $\Delta_G \leq 8 \ln(2nk/\eta)$ and the bound for the Laplacian is trivial when $\delta_G \leq 8 \ln(2nk/\eta)$. Moreover, the results of Lubetsky et al. [18] greatly improve upon Theorem 1.1 for $d$-regular Ramanujan graphs in the range $k \gg n^{3/2}$: in that case, the new eigenvalues are all $O(\sqrt{d \ln d})$.

One corollary of Theorem 1.1 is the following result.

**Corollary 1.1.** *In the setting of Theorem 1.1, let $k = \prod_{i=1}^{s}$ with $k_1, \ldots, k_s \in \mathbb{N}\backslash\{0,1\}$ and consider two different random graphs (see Definition 3.1 for the precise properties of a standard random $k$-lift):*

- *$G^{(k)}$ is a standard random $k$-lift of $G$: that is to say, each random matching $\mathcal{M}_{vw}$ appearing in the construction of $G^{(k)}$ is uniformly distributed over all matchings of $\{(v,i)\}_{i=1}^{k}$ and $\{(w,j)\}_{j=1}^{k}$, and the matchings are independent.*
- *$\tilde{G}^{(k)} = G_s$ where $G_0 = G$ and, for each $1 \leq i \leq s$, conditionally on $\{G_0, G_1, \ldots, G_{i-1}\}$, $G_i$ is a standard random $k_i$-lift of $G_{i-1}$.*

---

[1]It is possible to find connected examples with similar behavior.

Let $A^{(k)}$ and $\mathcal{L}^{(k)}$ denote the adjacency matrix and Laplacian of $G^{(k)}$ and define $\tilde{A}^{(k)}$ and $\tilde{\mathcal{L}}^{(k)}$ similarly. Then (with an appropriate labeling of the vertices of the two graphs):

$$\mathbb{P}\left(\|A^{(k)} - \tilde{A}^{(k)}\| \le 4\sqrt{2\Delta_G \ln(4nk/\eta)}\right) \ge 1 - \eta$$

and

$$\mathbb{P}\left(\|\mathcal{L}^{(k)} - \tilde{\mathcal{L}}^{(k)}\| \le 4\sqrt{\frac{2\ln(4nk/\eta)}{\delta_G}}\right) \ge 1 - \eta.$$

This is interesting because the distributions of $G^{(k)}$ and $\tilde{G}^{(k)}$ can be very different. For instance, let $k_1 = k_2 = \cdots = k_s = 2$. If $s$ is constant and the number of vertices is large enough, all $2^s!$ possible permutations will be seen in the matchings of $\{v\} \times [k]$ with $\{w\} \times [k]$ for $vw \in E$. On the other hand, only $2^s$ possible permutations will be seen in $\tilde{G}^{(k)}$.

Theorem 1.1 will be deduced from a concentration result for sums of independent random *matrices*. In what follows $\mathbb{C}_{\text{Herm}}^{d \times d}$ is the space of $d \times d$ Hermitian matrices with complex entries, the expectations of matrices are defined entrywise and $\|\cdot\|$ is the operator norm (see Section 2.2 and Section 2.4 for these and related definitions).

**Theorem 1.2** (Proven in Section 5). *Let $X_1, \ldots, X_m$ be independent random matrices, defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with values in $\mathbb{C}_{\text{Herm}}^{d \times d}$. Suppose that $\sigma^2 > 0$ is such that:*

$$\left\|\sum_{i=1}^{m} X_i^2\right\| \le \sigma^2 \text{ almost surely.}$$

*Then for all $t \ge 0$,*

$$\mathbb{P}\left(\left\|\sum_{i=1}^{m}(X_i - \mathbb{E}[X_i])\right\| \ge t\right) \le 2d\, e^{-\frac{t^2}{8\sigma^2}}.$$

Given this bound, Theorem 1.1 follows quite easily, while other proof techniques for bounding spectra of random matrices (such as the *trace method* [12, 10, 11, 17] and the discrepancy-based ideas of Feige and Ofek [9]) can be quite technical. Theorem 1.2 will be deduced in Section 5 from a technical Lemma in [20], which was used to give a new proof of a concentration bound by Rudelson. A slightly sharper variant of Theorem 1.2 could be deduced from a form of Freedman's inequality for matrix martingales that

was recently proved by the author and applied to general random graph models with independent edges [19]. More recently, Tropp [22] has proved stronger inequalities in the same spirit. One can show that Theorem 1.2 and the related results just discussed have much smaller "variance" terms than related results in [1, 7]; see [22], Section 3.1 in [20] and Remark 3.1 in [19] for details.

The remainder of the paper is organized as follows. After the preliminary Section 2, we collect some basic facts about $k$-lifts in Section 3. We prove the Theorem and its Corollary in Section 4. Section 5 contains the proof of Theorem 1.2. The last section presents some extensions and open questions.

## 2. Preliminaries

### 2.1. Basic notation

For a natural number $m \in \mathbb{N}\backslash\{0\}$, $[m]$ is the set of all integers $1 \leq i \leq m$.

We will frequently speak of *multisets* $S$. Given a ground set $\mathcal{S}$ (which will usually be $\mathbb{R}$), a multiset $S$ is defined by a function $m_S : \mathcal{S} \to \mathbb{N}$. Informally, we will think of $S$ as a set where each $x \in A$ appears $m_S(x)$ times and we will refer to this quantity as the *multiplicity* of $x$. We say that $x$ belongs to $S$ ($x \in S$) if $m_S(x) > 0$.

For two multisets $S_1, S_2$ over the same ground set $\mathcal{S}$ and with corresponding functions $m_{S_1}, m_{S_2}$, we say that $S_1 \subset S_2$ if for all $x \in \mathcal{S}$ $m_{S_1}(x) \leq m_{S_2}(x)$. The difference $S_2\backslash S_1$ is the multiset where each $x \in \mathcal{S}$ has multiplicity $\max\{m_{S_2}(x) - m_{S_1}(x), 0\}$.

### 2.2. Linear algebra

For given $d_r, d_c \in \mathbb{N}\backslash\{0\}$, $\mathbb{R}^{d_r \times d_c}$ (resp. $\mathbb{C}^{d_r \times d_c}$) is the space of $d_r \times d_c$ matrices with entries in $\mathbb{R}$ (resp. $\mathbb{C}$).

For $A \in \mathbb{R}^{d_r \times d_c}$, $A^\dagger \in \mathbb{R}^{d_c \times d_r}$ is the transpose of $A$; similarly, for $B \in \mathbb{C}^{d_r \times d_c}$, $B^* \in \mathbb{C}^{d_c \times d_r}$ is the conjugate transpose of $B$. We identify $\mathbb{R}^d$ and $\mathbb{C}^d$ with $\mathbb{R}^{d \times 1}$ and $\mathbb{C}^{d \times 1}$ (resp.), so that the standard inner product of $x, y \in \mathbb{R}^d$ is $x^\dagger y$.

$\mathbb{C}^{d \times d}_{\mathrm{Herm}}$ is the space of $d \times d$ *Hermitian matrices*, which are the $A \in \mathbb{C}^{d \times d}$ with $A^* = A$. Similarly, $\mathbb{R}^{d \times d}_{\mathrm{Sym}}$ is the space of all $d \times d$ real matrices that are symmetric in the sense that $A = A^\dagger$.

For a vector $v \in \mathbb{R}^d$ or $\mathbb{C}^d$, $\|v\|$ is its Euclidean norm. The *operator norm* of $A \in \mathbb{R}^{d \times d}$ is:

$$\|A\| \equiv \max_{v \in \mathbb{R}^d, \|v\|=1} \|Av\|.$$

Finally, the canonical basis vectors for $\mathbb{R}^d$ is denoted by $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_d$.

**2.2.1. The spectral theorem**   We recall the standard spectral theorem: for any $A \in \mathbb{R}^{d \times d}_{\mathrm{Sym}}$ there exists a set $S \subset \mathbb{R}$ and orthogonal projections $\{P_\alpha\}_{\alpha \in S}$ with orthogonal ranges such that:

$$\sum_{\alpha \in S} \alpha \, P_\alpha = A \text{ and } \sum_{\alpha \in S} P_\alpha = I_d,$$

where $I_d$ is the $d \times d$ identity matrix. The numbers $\alpha \in S$ are called the eigenvalues of $A$ and the vectors $v$ in the range of $P_\alpha$ are eigenvectors corresponding to a given $\alpha$. The spectrum of $A$, denoted by $\mathrm{spec}(A)$, is the multiset where each $\alpha \in S$ appears with multiplicity equal to the rank of $P_\alpha$.

One useful consequence of the spectral decomposition is that

$$\|A\| = \max_{\alpha \in \mathrm{spec}(A)} |\alpha|.$$

**2.2.2. Tensor products**   It will be convenient to represent the matrices of lifts via tensor products. The *tensor product* of $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$, denoted by $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$, is the set of all formal linear combinations of vectors of the form $\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2}$ with $1 \leq i_b \leq d_b$ for $b = 1, 2$. [We will abuse notation and assume that $e_i \in \mathbb{R}^{d_1} \cap \mathbb{R}^{d_2}$ for $i \leq \min\{d_1, d_2\}$.]

Similarly, if $v_b = \sum_{j_b=1}^{d_b} v_{b,j_b} \mathbf{e}_{j_b}$ $(b = 1, 2)$, the tensor product of $v_1 \otimes v_2$ is defined by the "distributive rule":

$$v_1 \otimes v_2 \equiv \sum_{j_1=1}^{d_1} \sum_{j_2=1}^{d_2} v_{1,j_1} v_{2,j_2} \mathbf{e}_{j_1} \otimes \mathbf{e}_{j_2}.$$

There exists a unique inner product on $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$, denoted by $(\cdot, \cdot\cdot)$, such that for all $v_1, w_1 \in \mathbb{R}^{d_1}$ and $v_2, w_2 \in \mathbb{R}^{d_2}$,

$$(v_1 \otimes v_2, w_1 \otimes w_2) = (v_1^\dagger w_1)\,(v_2^\dagger w_2).$$

Moreover, the tensor product of $A_1 \in \mathbb{R}^{d_1 \times d_1}$ and $A_2 \in \mathbb{R}^{d_2 \times d_2}$ is the unique linear operator $A_1 \otimes A_2$ from $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$ to itself that satisfies:

$$\forall 1 \leq i_1 \leq d_1, \, \forall 1 \leq i_2 \leq d_2, \, (A_1 \otimes A_2)(\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2}) = (A_1 \mathbf{e}_1) \otimes (A_2 \mathbf{e}_2).$$

This definition implies that:

$$(2.1) \qquad \forall u, v \in \mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}, \, (u, (A_1 \otimes A_2)v) = ((A_1^\dagger \otimes A_2^\dagger)u, v).$$

i.e. $A_1^\dagger \otimes A_2^\dagger$ is the adjoint of $A_1 \otimes A_2$. In particular, for all $A_1 \in \mathbb{R}_{\mathrm{Sym}}^{d_1 \times d_1}$ and $A_2 \in \mathbb{R}_{\mathrm{Sym}}^{d_2 \times d_2}$, $A_1 \otimes A_2$ is self-adjoint in the sense that:

$$\forall u, v \in \mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}, \ (u, (A_1 \otimes A_2)v) = ((A_1 \otimes A_2)u, v).$$

Notice that $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$ is isomorphic to $\mathbb{R}^{d_1 d_2}$, in the sense that any bijection $\psi : [d_1] \times [d_2] \to [d_1 d_2]$ can be "lifted" to an invertible, inner-product-preserving linear map:

$$\Psi : \mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2} \to \mathbb{R}^{d_1 d_2}$$

defined by the rule $\Psi(\mathbf{e}_i \otimes \mathbf{e}_j) = \mathbf{e}_{\psi(i,j)}$, $(i, j) \in [d_1] \times [d_2]$. Under this map, self-adjoint maps over $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$ correspond to symmetric matrices over $\mathbb{R}^{d_1 d_2}$ and vice versa. Therefore, one may also state a spectral theorem over $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$; we omit the details.

## 2.3. Concepts from graph theory

For our purposes a graph $G = (V, E)$ consists of a finite set $V$ of vertices and a set $E$ of edges, which are subsets of size 2 of $V$. Unless otherwise noted, we will assume that $V = [n]$ for some integer $n \geq 2$, where $[n] \equiv \{1, 2, \ldots, n\}$. We will write edges as unordered pairs $vw$ or $\{v, w\}$ and make no distinction between $vw$ and $wv$. The *degree* $\mathrm{d}_G(v)$ of a vertex $v$ is the number of $w \in V \backslash \{v\}$ such that $vw \in E$.

Assume for simplicity that $V = [n]$, or more generally, that the elements of $V$ are labeled $v_1, \ldots, v_n$. The *adjacency matrix* of $G$ is the $n \times n$ matrix $A \in \mathbb{R}_{\mathrm{Sym}}^{n \times n}$ with zeros on the diagonal and such that, for all $1 \leq i < j \leq n$, the $(i, j)$-th entry of $A$ is 1 if $v_i v_j \in E$ and 0 otherwise. When $V = [n]$, this reads:

$$(2.2) \qquad A \equiv \sum_{ij \in E} (\mathbf{e}_i \mathbf{e}_j^\dagger + \mathbf{e}_j \mathbf{e}_i^\dagger).$$

The *Laplacian* $\mathcal{L}$ of $G$ is the matrix:

$$\mathcal{L} = I_n - D^{-1/2} A D^{-1/2}$$

where $D$ is the $n \times n$ diagonal matrix whose $(i, i)$-th entry is $\mathrm{d}_G(v_i)$ if $\mathrm{d}_G(v_i) \neq 0$ (if $\mathrm{d}_G(v_i) = 0$ we may define the corresponding entry arbitrarily).

If all degrees are non-zero, one can write this as follows:

$$(2.3) \qquad \mathcal{L} = I_n - \sum_{ij \in E} \frac{(\mathbf{e}_i \mathbf{e}_j^\dagger + \mathbf{e}_j \mathbf{e}_i^\dagger)}{\sqrt{\mathrm{d}_G(v_i) \mathrm{d}_G(v_j)}}.$$

**Remark 2.1.** The normalized Laplacian will be simply called *the* Laplacian in this paper. Many readers will be familiar with the related definition of the so-called *combinatorial Laplacian* $L = D - A$, but the normalized Laplacian turns out to be better behaved. See [13] for details.

## 2.4. Probability with matrices

We will be dealing with random matrices (and random linear operators) throughout the paper. Following common practice, we will always assume that we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ in the background where all random variables are defined.

Call a map $X : \Omega \to \mathbb{C}^{d \times d}_{\mathrm{Herm}}$ a random $d \times d$ Hermitian matrix (or a $\mathbb{C}^{d \times d}_{\mathrm{Herm}}$-valued random variable) if for each $1 \leq i, j \leq n$, the function $X(i, j) : \Omega \to \mathbb{C}^{d \times d}$ corresponding to the $(i, j)$-th entry of $X$ is $\mathcal{F}$-measurable, or equivalently, if for each Borel subset $S \subset \mathbb{C}^{d \times d}_{\mathrm{Herm}}$, $X^{-1}(S) \in \mathcal{F}$. We say that $X$ is integrable if all the entries of $X$ are integrable. If $X$ is integrable, one defines $\mathbb{E}[X]$ entrywise: the $(i, j)$th entry of $\mathbb{E}[X]$ is $\mathbb{E}[X(i, j)]$. We will also use analogous definitions for $X : \Omega \to \mathbb{R}^{d \times d}_{\mathrm{Sym}}$. We will essentially ignore all measurability and integrability issues in the remainder of the paper, as these can be dealt with in a rather straightforward manner.

One can easily check that if $X$ is a random integrable $d \times d$ Hermitian matrix and $A \in \mathbb{C}^{d \times d}_{\mathrm{Herm}}$ is deterministic, $\mathbb{E}[AX] = A\mathbb{E}[X]$.

## 3. Lifts of graphs

Our goal here is to review the construction of lifts of graphs outlined in the introduction and to prove some elementary facts that will be useful later on. Other perspectives on these objects can be found in [2].

Recall that a *matching* of two finite, disjoint, non-empty sets $A, B$ is a set of pairs:

$$\mathcal{M} = \{\{a_i, b_i\} : i = 1, \ldots, m\}$$

where $(a_1, \ldots, a_m)$ is a permutation of the elements of $A$ and $(b_1, \ldots, b_m)$ is a permutation of the elements of $B$. Notice that the existence of a matching $\mathcal{M}$ as above implies that $|A| = |B| = m$.

Now let $G$ be a graph with vertex set $V = [n]$ and edge set $E$. Given $k \in \mathbb{N}\backslash\{0,1\}$, a $k$-lift $G^{(k)}$ of $G$ is determined by a choice of matchings:

$$\{\mathcal{M}_{ij} : ij \in E\},$$

where for each $ij \in E$, $\mathcal{M}_{ij}$ is a matching of $\{i\} \times [k]$ and $\{j\} \times [k]$. $G^{(k)}$ is the graph with vertex set $[n] \times [k]$ and edge set $E^{(k)} = \cup_{ij \in E} \mathcal{M}_{ij}$.

We will also need the following definition.

**Definition 3.1.** Let $G = (V, E)$ be a graph. A *weakly random $k$-lift* of $G$ is a $k$-lift $G^{(k)}$ of $G$ as above where the matchings $\mathcal{M}_{ij}$ are independently chosen and satisfy:

$$\forall ij \in E, \ \forall (\ell, r) \in [k]^2 \ : \ \mathbb{P}\left(\{(i, \ell), (j, r)\} \in \mathcal{M}_{vw}\right) = \frac{1}{k}.$$

If, in addition, each $\mathcal{M}_{ij}$ is uniformly distributed over all matchings of $\{i\} \times [k]$ and $\{j\} \times [k]$, $G^{(k)}$ is a *standard random $k$-lift* of $G$.

### 3.1. Graph matrices and tensor products

It is convenient to represent the matrices corresponding to $G^{(k)}$ in the tensor space $\mathbb{R}^n \otimes \mathbb{R}^k$. That is to say, we will write down a linear operator $A^{(k)}$ over $\mathbb{R}^n \otimes \mathbb{R}^k$ such that for all $(i, \ell), (j, r) \in [n] \times [k]$,

$$(\mathbf{e}_i \otimes \mathbf{e}_\ell, A^{(k)}(\mathbf{e}_j \otimes \mathbf{e}_r)) = \begin{cases} 1 & \text{if } \{(i,\ell),(j,r)\} \in E^{(k)}; \\ 0 & \text{otherwise.} \end{cases}$$

This is satisfied by:

$$A^{(k)} = \sum_{\{(i,\ell),(j,r)\} \in E^{(k)}} (\mathbf{e}_i \mathbf{e}_j^\dagger) \otimes (\mathbf{e}_\ell \mathbf{e}_r^\dagger) + (\mathbf{e}_j \mathbf{e}_i^\dagger) \otimes (\mathbf{e}_r \mathbf{e}_\ell^\dagger).$$

Another way of writing $A^{(k)}$ will be more useful later on. Defining:

$$(3.1) \qquad V_{(i,j)} \equiv \sum_{(\ell,r) \in [k]^2 \, : \, \{(i,\ell),(j,r)\} \in \mathcal{M}_{ij}} \mathbf{e}_\ell \mathbf{e}_r^\dagger, \ (ij \in E),$$

we have:

$$(3.2) \qquad A^{(k)} = \sum_{ij \in E} \mathbf{e}_i \mathbf{e}_j^\dagger \otimes V_{(i,j)} + \mathbf{e}_j \mathbf{e}_i^\dagger \otimes V_{(j,i)}.$$

We emphasize that the definition of $V_{(i,j)}$ is *not* symmetric with respect to $i, j$: in fact, a simple computation shows that $V_{(j,i)} = V_{(i,j)}^{\dagger} = V_{(i,j)}^{-1}$.

The Laplacian $\mathcal{L}^{(k)}$ of $G^{(k)}$ can be similarly written as a linear operator over $\mathbb{R}^n \otimes \mathbb{R}^k$. The key point to notice is that all copies of $i \in [n]$ in $G^{(k)}$ have the same degree, i.e.:

$$\forall \ell \in [k], \, \mathrm{d}_{G^{(k)}}((i,\ell)) = \mathrm{d}_G(i).$$

A simple calculation (omitted) shows that:

$$(3.3) \qquad \mathcal{L}^{(k)} = I_n \otimes I_k - \sum_{ij \in E} \frac{\mathbf{e}_i \mathbf{e}_j^{\dagger} \otimes V_{(i,j)} + \mathbf{e}_j \mathbf{e}_i^{\dagger} \otimes V_{(j,i)}}{\sqrt{\mathrm{d}_G(i)\mathrm{d}_G(j)}}.$$

### 3.2. Old and new eigenvalues

We now draw a connection between the spectrum and eigenvalues of $A$ and $A^{(k)}$. All arguments here also appeared on previous papers on graph lifts (e.g. [5, 10]).

**Proposition 3.1.** *The spectrum of the adjacency matrix $A$ of $G$ is contained in the spectrum of $A^{(k)}$ (counting multiplicities). Moreover, if*

$$\mathrm{new}(A^{(k)}) \equiv \mathrm{spec}(A^{(k)}) \backslash \mathrm{spec}(A)$$

*is the difference of the two spectra as multisets,*

$$\max_{\eta \in \mathrm{new}(A^{(k)})} |\eta| = \|A^{(k)} - A \otimes \Pi_k\|$$

*where $\Pi_k$ is the $k \times k$ matrix with all entries equal to $1/k$.*

Essentially the same argument shows a related result for the Laplacian $\mathcal{L}^{(k)}$ of $G^{(k)}$ (proof omitted).

**Proposition 3.2.** *The spectrum of the Laplacian $\mathcal{L}$ of $G$ is contained in the spectrum of $\mathcal{L}^{(k)}$ (counting multiplicities). Moreover, if*

$$\mathrm{new}(\mathcal{L}^{(k)}) \equiv \mathrm{spec}(\mathcal{L}^{(k)}) \backslash \mathrm{spec}(\mathcal{L})$$

*is the difference of the two spectra as multisets,*

$$\max_{\eta \in \mathrm{new}(\mathcal{L}^{(k)})} |1 - \eta| = \|\mathcal{L}^{(k)} - (I_n \otimes I_k - (I - \mathcal{L}) \otimes \Pi_k)\|$$

*where $\Pi_k$ is the $k \times k$ matrix with all entries equal to $1/k$.*

*Proof of Proposition 3.1.* Let $\mathbf{1}_k \in \mathbb{R}^k$ be the vector with all coordinates equal to 1. Notice that $V_{(i,j)}\mathbf{1}_k = \Pi_k\mathbf{1}_k = \mathbf{1}_k$ for all $i,j$ with $ij \in E$. Therefore, for all vectors $v \in \mathbb{R}^n$,

$$A^{(k)}(v \otimes \mathbf{1}_k) = (Av) \otimes \mathbf{1}_k = (A \otimes \Pi_k)\mathbf{1}_k.$$

In particular, if $v$ is an eigenvector of $A$ with eigenvalue $\lambda$, $v \otimes \mathbf{1}_k$ is an eigenvector of both $A^{(k)}$ and $A \otimes \Pi_k$, with the same eigenvalue $\lambda$ for both matrices. It follows that each eigenvalue $\lambda$ of $A$ with multiplicity $m$ is an eigenvalue of $A^{(k)}$ with multiplicity $\geq m$, which is the first assertion in the Proposition.

Any new eigenvalue $\eta \in \text{new}(A^{(k)})$ must correspond an eigenvector $w \in \mathbb{R}^n \otimes \mathbb{R}^k$ that is orthogonal to $v \otimes \mathbf{1}_k$ for all eigenvectors $v$ of $A$ corresponding to "old" eigenvalues. Since the eigenvectors of $A$ span $\mathbb{R}^n$, any $w$ as above must be orthogonal to the subspace:

$$H \equiv \{v \otimes \mathbf{1}_k \,:\, v \in \mathbb{R}^n\} \subset \mathbb{R}^n \otimes \mathbb{R}^k.$$

In particular,

$$\max_{\eta \in \text{new}(A^{(k)})} |\eta| = \max_{w \in H^\perp} \|A^{(k)}w\|.$$

To finish, we must show that the RHS equals $\|A^{(k)} - A \otimes \Pi_k\|$. We have already seen that the operators $A^{(k)}$ and $A \otimes \Pi_k$ have $H$ as an invariant subspace and that their restrictions to that subspace are equal. This implies that $H^\perp$ must also be invariant and moreover:

$$\|A^{(k)} - A \otimes \Pi_k\| = \max_{w \in H^\perp \,:\, \|w\|=1} \|A^{(k)}w - (A \otimes \Pi_k)w\|.$$

Now notice that:

$$H^\perp \equiv \text{span}\{x \otimes y \,:\, x \in \mathbb{R}^n, y \in \mathbb{R}^k, \, y \perp \mathbf{1}_k\}.$$

Moreover, for all $x \otimes y$ as above,

$$(A \otimes \Pi_k)(x \otimes y) = (Ax) \otimes (\Pi_k y) = 0$$

since $\Pi_k$ is the projection onto the line spanned by $\mathbf{1}_k$. By linearity, this implies that $(A \otimes \Pi_k)w = 0$ for all $w \in H^\perp$, which results in the desired equality:

$$\|A^{(k)} - A \otimes \Pi_k\| = \max_{w \in H^\perp \,:\, \|w\|=1} \|A^{(k)}w\|.$$

$\square$

# 4. Main proofs

Propositions 3.1 and 3.2 show that in order to prove Theorem 1.1, one must bound the difference between certain matrices. We attack this problem from the perspective of concentration of measure. As it turns out, $A \otimes \Pi_k$ is the expected value of $A^{(k)}$ and $I_n \otimes I_k - (I_n - \mathcal{L}) \otimes \Pi_k$ is the expected value of $\mathcal{L}^{(k)}$. The concentration inequality in Theorem 1.2 will ensure that $A^{(k)}$ and $\mathcal{L}^{(k)}$ are likely to be close to their respective expected values. Once this is achieved, Theorem 1.1 and its Corollary will easily follow.

## 4.1. Proof of the main theorem

In this section we prove Theorem 1.1.

*Proof of Theorem 1.1.* We start with the result for the adjacency matrix. Proposition 3.1 implies that it is necessary and sufficient to prove that:

$$(4.1) \qquad [\mathbf{Goal}] \ \mathbb{P}\left( \|A^{(k)} - A \otimes \Pi_k\| \leq 2\sqrt{2\Delta_G \ln(2nk/\eta)} \right) \geq 1 - \eta.$$

We will restate this as a concentration bound for the sum of random matrices. Recall from Section 3.1 that:

$$A^{(k)} = \sum_{ij \in E} Z_{ij} \text{ where } Z_{ij} = \mathbf{e}_j \mathbf{e}_i^\dagger \otimes V_{(i,j)} + \mathbf{e}_i \mathbf{e}_j^\dagger \otimes V_{(j,i)}.$$

We will later apply Theorem 1.2 to $\sum_{ij \in E} Z_{ij}$. For this we will need the following claim.

**Claim 4.1.** *The $Z_{ij}$ are self-adjoint and independent.*

*Proof of Claim 4.1.* The first assertion comes from (2.1) and the fact that $V_{(i,j)}^\dagger = V_{(j,i)}$ (cf. Section 3.1). To prove the second assertion, notice that each $Z_{ij}$ is determined by the random matching $\mathcal{M}_{ij}$. Since the matchings are independent according to Definition 3.1, the $Z_{ij}$ are also independent. (Incidentally, this is the key point where the independent of the matchings is used.) $\qquad\square$

The next claim relates the Theorem 1.2 to (4.1).

**Claim 4.2.** $\sum_{ij \in E} \mathbb{E}\left[ Z_{ij} \right] = A \otimes \Pi_k.$

*Proof of Claim 4.2.* We first compute $\mathbb{E}[Z_{ij}]$ for a fixed $ij \in E$. It is not hard to show that this is:

$$\mathbb{E}[Z_{ij}] = \mathbf{e}_j \mathbf{e}_i^\dagger \otimes \mathbb{E}\left[V_{(i,j)}\right] + \mathbf{e}_i \mathbf{e}_j^\dagger \otimes \mathbb{E}\left[V_{(j,i)}\right].$$

The $(\ell, r)$-th entry of $V_{(i,j)}$ is an indicator random variable that is equal to 1 iff $(i, \ell)$ and $(j, r)$ are connected in the matching. By the definition of a weakly random $k$-lift (cf. Definition 3.1), this happens with probability $1/k$, therefore each entry of $V_{(i,j)}$ has expected value $1/k$. This implies that $\mathbb{E}[V_{(i,j)}]$ is precisely the matrix $\Pi_k$ in the Theorem. Similarly, $\mathbb{E}[V_{(j,i)}] = \Pi_k$. We deduce that:

(4.2) $$\mathbb{E}[Z_{ij}] = (\mathbf{e}_j \mathbf{e}_i^\dagger + \mathbf{e}_i \mathbf{e}_j^\dagger) \otimes \Pi_k.$$

We deduce:

$$\sum_{ij \in E} \mathbb{E}[Z_{ij}] = \left(\sum_{ij \in E} \mathbf{e}_j \mathbf{e}_i^\dagger + \mathbf{e}_i \mathbf{e}_j^\dagger\right) \otimes \Pi_k. = A \otimes \Pi_k.$$

$\square$

The two claims show that

(4.3) $$A^{(k)} - A \otimes \Pi_k = \sum_{ij \in E} (Z_{ij} - \mathbb{E}[Z_{ij}]),$$

where the RHS is a centered sum of independent, self-adjoint random linear operators. One may recall from Section 2.2.2 that self-adjoint linear operators over $\mathbb{R}^n \otimes \mathbb{R}^k$ correspond to symmetric matrices over $\mathbb{R}^{nk}$; therefore, we can apply Theorem 1.2 with $d = nk$ to the sum once we compute the variance parameter $\sigma^2$.

We start by computing $Z_{ij}^2$ for a fixed $ij \in E$. One can check that:

$$Z_{ij}^2 = \mathbf{e}_i \mathbf{e}_i^\dagger \otimes (V_{(j,i)} V_{(i,j)}) + \mathbf{e}_j \mathbf{e}_j^\dagger \otimes (V_{(i,j)} V_{(j,i)}).$$

Now recall from Section 3.1 that $V_{(j,i)} = V_{(i,j)}^{-1}$ and deduce that:

$$Z_{ij}^2 = (\mathbf{e}_i \mathbf{e}_i^\dagger + \mathbf{e}_j \mathbf{e}_j^\dagger) \otimes I_k.$$

Summing up those terms, we arrive at:

$$\sum_{ij \in E} Z_{ij}^2 = \left[\sum_{ij \in E} (\mathbf{e}_i \mathbf{e}_i^\dagger + \mathbf{e}_j \mathbf{e}_j^\dagger)\right] \otimes I_k = \left[\sum_{i=1}^n \mathrm{d}_G(i) \, \mathbf{e}_i \mathbf{e}_i^\dagger\right] \otimes I_k.$$

Given a symmetric matrix $B$, the eigenvalues of $B \otimes I_k$ are precisely the eigenvalues of $B$. It follows that:

$$\left\| \sum_{ij \in E} Z_{ij}^2 \right\| = \left\| \sum_{i=1}^{n} d_G(i) \, \mathbf{e}_i \mathbf{e}_i^\dagger \right\|.$$

But the matrix on the RHS is diagonal with non-negative entries, hence its largest eigenvalue is the largest entry on the diagonal, which is $\max_i d_G(i) = \Delta_G$. We deduce that one may take $\sigma^2 = \Delta_G$ and $d = nk$ (as already seen) in Theorem 1.2 to deduce:

$$\mathbb{P}\left( \|A^{(k)} - A \otimes \Pi_k\| \geq t \right) \leq 2nk \, e^{-\frac{t^2}{8\Delta_G^2}}.$$

Taking $t \equiv 2\sqrt{2\Delta_G \ln(2nk/\eta)}$ makes the RHS smaller than $\eta$. This proves (4.1), which (as seen above) is equivalent to the desired assertion via Proposition 3.1.

The proof for the Laplacian is quite similar and we will present it in less detail. Define $R \equiv (I_n \otimes I_k - (I - \mathcal{L}) \otimes \Pi_k)$. We use Proposition 3.2 in order to restate the desired inequality as:

$$(4.4) \qquad [\textbf{Goal}] \ \ \mathbb{P}\left( \|\mathcal{L}^{(k)} - R\| \leq 2\sqrt{\frac{2 \ln(2nk/\eta)}{\delta_G}} \right) \geq 1 - \eta,$$

Using equations (2.3) and (3.3), we see that:

$$R - \mathcal{L}^{(k)} = \sum_{ij \in E} \frac{\mathbf{e}_j \mathbf{e}_i^\dagger \otimes (V_{(i,j)} - \Pi_k) + \mathbf{e}_i \mathbf{e}_j^\dagger \otimes (V_{(j,i)} - \Pi_k)}{\sqrt{d_G(i) d_G(j)}}$$

$$(4.5) \qquad = \sum_{ij \in E} \left( \frac{Z_{ij}}{\sqrt{d_G(i) d_G(j)}} - \mathbb{E}\left[ \frac{Z_{ij}}{\sqrt{d_G(i) d_G(j)}} \right] \right)$$

with the same $Z_{ij}$ from the first part (we used here the formula for $\mathbb{E}[Z_{ij}]$ appearing in the proof of Claim 4.2).

In analogy with Claim 4.1, we note that the random matrices

$$\{Z_{ij}/\sqrt{d_G(i) d_G(j)}\}_{ij \in E}$$

are independent and self-adjoint. Thus we may apply Theorem 1.2 with $d = nk$ to the sum in (4.5) once we find $\sigma^2$. For this we follow the calculations

in the first part of the proof to deduce:

$$\sum_{ij\in E}\left(\frac{Z_{ij}}{\sqrt{\mathrm{d}_G(i)\mathrm{d}_G(j)}}\right)^2 = \left\{\sum_{ij\in E}\frac{\mathbf{e}_i\mathbf{e}_i^\dagger + \mathbf{e}_j\mathbf{e}_j^\dagger}{\mathrm{d}_G(i)\mathrm{d}_G(j)}\right\}\otimes I_k$$

$$= \left\{\sum_{i=1}^{n}\left(\sum_{j\,:\,ij\in E}\frac{1}{\mathrm{d}_G(i)\mathrm{d}_G(j)}\right)\mathbf{e}_i\mathbf{e}_i^\dagger\right\}\otimes I_k.$$

The largest eigenvalue of this matrix is simply the largest eigenvalue of the first term of the tensor product in the RHS, which is:

$$\max_{i\in[n]}\sum_{j\,:\,ij\in E}\frac{1}{\mathrm{d}_G(i)\mathrm{d}_G(j)} \le \frac{1}{\delta_G}\max_{i\in[n]}\sum_{j\,:\,ij\in E}\frac{1}{\mathrm{d}_G(i)} = \frac{1}{\delta_G}.$$

Therefore, we may take $\sigma^2 = \delta_G^{-1}$ in Theorem 1.2 and deduce:

$$\mathbb{P}\left(\|\mathcal{L}^{(k)} - (I_n\otimes I_k - (I-\mathcal{L})\otimes\Pi_k)\| \ge t\right) \le 2nk\,e^{-\frac{t^2\delta_G}{8}}.$$

Taking:

$$t \equiv 2\sqrt{\frac{2\ln(2nk/\eta)}{\delta_G}}$$

makes the RHS $\le \eta$ and finishes the proof of (4.4), which is equivalent to the second assertion in the Theorem (cf. Proposition 3.2). $\square$

## 4.2. Proof of the corollary

*Proof of Corollary 1.1.* We only present the proof for the adjacency matrix result; the argument for the Laplacian is exactly the same.

The adjacency matrix $A^{(k)}$ of the graph $G^{(k)}$ satisfies:

$$(4.6)\qquad \mathbb{P}\left(\|A^{(k)} - A\otimes\Pi^{(k)}\| \le 2\sqrt{2\Delta_G\ln(4nk/\eta)}\right) \ge 1 - \frac{\eta}{2}.$$

This is precisely what we showed in the course of the proof of Theorem 1.1, cf. (4.1). It also follows from combining the Theorem with Proposition 3.1.

We claim that the same bound holds for $\tilde{A}^{(k)}$, after a suitable relabeling of the vertices. The vertex set of the graph $\tilde{G}^{(k)}$ is $[n]\times K$ where

$$K = [k_1]\times[k_2]\times\cdots\times[k_s].$$

A simple induction argument shows that $\tilde{G}^{(k)}$ is also a lift of $G$, in the sense that its edge set $\tilde{E}^{(k)}$ is a union:

$$\tilde{E}^{(k)} = \bigcup_{ij \in E} \tilde{\mathcal{M}}_{ij},$$

where $\tilde{\mathcal{M}}_{ij}$ is a matching of $\{i\} \times K$ and $\{j\} \times K$.

It is easy to see that these matchings are independent, because they correspond to successive matchings of the lifted images of distinct edges of $G$. Moreover, two vertices $(i, \ell_1, \ldots, \ell_s) \in \{i\} \times K$ and $(j, r_1, \ldots, r_s) \in j \times \{j\} \times K$ are matched in $\tilde{\mathcal{M}}_{ij}$ if $(i, \ell_1)$ is matched to $(j, r_1)$ in $G_1$ *and* $(i, r_1, r_2)$ is matched to $(j, r_1, r_2)$ in $G_2$ *and* ... $(i, \ell_1, \ldots, \ell_s)$ is matched to $(j, r_1, \ldots, r_s)$ in $G_s$. The recipe for constructing $G_s$ implies that the probability of this event is:

$$\mathbb{P}\left(\{(i, \ell_1, \ldots, \ell_s), (j, r_1, \ldots, r_s)\} \in \tilde{\mathcal{M}}_{ij}\right) = \frac{1}{k_1 k_2 \ldots k_s} = \frac{1}{k}.$$

Thus if we label the elements of $K$ with the numbers $1, 2, \ldots, k$, we see that $\tilde{G}^{(k)}$ satisfies the assumptions of the Theorem. It follows that, just as in the case of $G^{(k)}$,

$$\mathbb{P}\left(\|\tilde{A}^{(k)} - A \otimes \Pi^{(k)}\| \leq 2\sqrt{2\Delta_G \ln(4nk/\eta)}\right) \geq 1 - \frac{\eta}{2}.$$

Putting this together with (4.6) finishes the proof.  $\square$

## 5. Proof of the concentration inequality for matrices

In this section we prove Theorem 1.2. Our proof is based on results from [20] and we refer to that paper for further details on the basic results needed, such as the spectral mapping property.

Given $U \in \mathbb{C}_{\text{Herm}}^{d \times d}$, we let $\lambda_{\max}(U)$ denote the largest eigenvalue of $U$. Notice that:

$$\|U\| = \max\{\lambda_{\max}(U), \lambda_{\max}(-U)\},$$

hence:

$$\mathbb{P}\left(\|\sum_{i \leq n}(X_i - \mathbb{E}[X_i])\| \geq t\right) \leq \mathbb{P}\left(\lambda_{\max}\left[\sum_{i \leq n}(X_i - \mathbb{E}[X_i])\right] \geq t\right)$$

$$\text{(5.1)} \qquad\qquad\qquad + \mathbb{P}\left(\lambda_{\max}\left[\sum_{i \leq n}(\mathbb{E}[X_i] - X_i)\right] \geq t\right).$$

We will prove that the first term in the RHS is $\leq d\,e^{-t^2/4\sigma^2}$; the second term can be similarly controlled if we replace $X_i$ with $-X_i$.

To control the first term in the RHS of (5.1), we begin with the usual Bernstein trick, valid for any $s > 0$:

$$\mathbb{P}\left(\lambda_{\max}\left[\sum_{i\leq n}(X_i - \mathbb{E}\,[X_i])\right] \geq t\right)$$

(5.2)
$$\leq e^{-st}\,\mathbb{E}\left[\exp\left(s\lambda_{\max}\left[\sum_{i\leq n}(X_i - \mathbb{E}\,[X_i])\right]\right)\right].$$

Notice that $\lambda_{\max}$ defines a convex map over $\mathbb{C}^{d\times d}_{\mathrm{Herm}}$, because of the variational characterization

$$\lambda_{\max}(U) = \max_{|v|=1} v^*Uv.$$

A standard symmetrization argument (see e.g. Lemma 6.3 in Chapter 6 of [15]) implies:

(5.3)
$$\mathbb{E}\left[e^{s\lambda_{\max}\left(\sum_{i\leq n}(X_i - \mathbb{E}[X_i])\right)}\right] \leq \mathbb{E}\left[e^{2s\lambda_{\max}\left[\sum_{i\leq n}\epsilon_i\,X_i\right]}\right]$$

where the $\{\epsilon_i\}_{i=1}^n$ are iid unbiased over $\{-1,+1\}$ and independent from $\{X_j\}_{j=1}^n$. The spectral mapping property of matrix exponentials implies that we may move $\lambda_{\max}$ to outside the exponential.

$$\mathbb{E}\left[\exp\left(s\lambda_{\max}\left[\sum_{i\leq n}\epsilon_i\,X_i\right]\right)\right] = \mathbb{E}\left[\lambda_{\max}\left[\exp\left(s\sum_{i\leq n}\epsilon_i\,X_i\right)\right]\right].$$

Since a matrix exponential has non-negative eigenvalues and "trace = sum of eigenvalues", this gives:

(5.4)
$$\mathbb{E}\left[\exp\left(s\lambda_{\max}\left[\sum_{i\leq n}\epsilon_i\,X_i\right]\right)\right] \leq \mathbb{E}\left[\mathrm{Tr}\left[e^{s\sum_{i\leq n}\epsilon_i\,X_i}\right]\right],$$

where Tr denotes trace. We now employ the following key inequality from [20]:

**Lemma 5.1** (Lemma 2 in [20]). *If $\{\epsilon_i\}_{i=1}^n$ are as above, then for all $s \in \mathbb{R}$, $d \in \mathbb{N}\backslash\{0\}$ and all deterministic $A_1,\ldots,A_n \in \mathbb{C}^{d\times d}_{\mathrm{Herm}}$:*

$$\mathbb{E}\left[\mathrm{Tr}\left[e^{s\sum_{i=1}^n\epsilon_i A_i}\right]\right] \leq \mathrm{Tr}\left[e^{\frac{s^2\sum_{i=1}^n A_i^2}{2}}\right].$$

In our setting, we may condition on $2X_i = A_i$, $1 \leq i \leq n$, and then integrate the $\epsilon_i$ to deduce:

$$\mathbb{E}\left[\text{Tr}\left[e^{s\sum_{i\leq n}\epsilon_i X_i}\right]\right] \leq \mathbb{E}\left[\text{Tr}\left[e^{2s^2\sum_{i=1}^n X_i^2}\right]\right].$$

Now the assumption that $\|\sum_{i=1}^n X_i^2\| \leq \sigma^2$ a.s. implies that the $d$ eigenvalues of the exponential in the RHS lie between 1 and $e^{2s^2\sigma^2}$. In particular, the trace in the RHS is $\leq d\,e^{2s^2\sigma^2}$ almost surely. It follows from (5.4) that:

$$\mathbb{E}\left[\exp\left(s\lambda_{\max}\left[\sum_{i\leq n}\epsilon_i X_i\right]\right)\right] \leq d\,e^{2s^2\sigma^2}.$$

Combining this with (5.3) and (5.2) gives:

$$\forall s > 0,\ \mathbb{P}\left(\lambda_{\max}\left[\sum_{i\leq n}(X_i - \mathbb{E}[X_i])\right] \geq t\right) \leq d\,e^{-st+2\sigma^2 s^2}$$

and the Theorem follows from the choice $s = t/4\sigma^2$.

## 6. Extensions and open questions

*Lifts of Marked chains.* The argument we showed can be applied to lifts of weighted graphs, or equivalently, of *reversible Marked chains.* Let $P$ be the transition matrix of an irreducible Marked chain on $[n]$ that is reversible with respect to a probability measure $\pi$, meaning that $\pi(i)P(i,j) = \pi(j)P(j,i)$ for all $1 \leq i,j \leq n$. [This implies that $P$ has $n$ real eigenvalues.]

Choose a matching $\mathcal{M}_{ij}$ for each pair $1 \leq i \leq j \leq n$ in the same way as in Theorem 1.1 and consider a Marked chain $P^{(k)}$ on $[n] \times [k]$ with transition probabilities given by:

$$P^{(k)}((i,r),(j,\ell)) = \begin{cases} P(i,j) & \{(i,r),(j,\ell)\} \in \mathcal{M}_{ij}; \\ 0 & \text{if not.} \end{cases}$$

One can show (proof omitted) that the spectrum of $P^{(k)}$ contains that of $P$ and that all new eigenvalues of $P^{(k)}$ satisfy:

$$\mathbb{P}\left(\max_{\eta\in\text{new}(P^{(k)})}|\eta| \leq 16\sqrt{c_P\ln(nk/\eta)}\right) \geq 1 - \eta,$$

where

$$c_P \equiv \max_{i \in [n]} \sum_{j=1}^{n} \frac{\pi(j)P(j,i)^2}{\pi(i)}.$$

To prove this, one only needs to consider the symmetric matrix $Q$ with entries equal to

$$Q(i,j) \equiv \sqrt{\frac{\pi(i)}{\pi(j)}} \, P(i,j)$$

(which has the same spectrum as $P$) and the corresponding matrix $Q^{(k)}$ for the lifted chain $P^{(k)}$, which is reversible with respect to the probability distribution:

$$\pi^{(k)}(i,\ell) = \pi(i)/k \;\; ((i,\ell) \in [n] \times [k]).$$

Notice that the parameter $c_P$ always satisfies:

$$c_P \leq \max_{i \in [n]} \left\{ (\max_{r \in [n]} P(r,i)) \sum_{j=1}^{n} \frac{\pi(j)P(j,i)}{\pi(i)} \right\} = \max_{(i,r) \in [n]^2} P(r,i).$$

*Sharpness of the bound.* We do not know if the bound in Theorem 1.1 can be improved. For instance, could it be the case that all new eigenvalues of the adjacency matrix are $O\left(\sqrt{\Delta_G}\right)$ with high probability, at least when the minimum degree is $\Omega(\ln n)$? This would be similar to the Erdös-Rényi random graph [9] and also related to results on random regular graphs [11, 6]. An analysis of the proof of Theorem 1.1 shows that the only obstacle to obtaining such a bound is the $d$ term in Theorem 1.2, but that term is known to be necessary in general [19]. However, it might be possible to obtain better bounds in the graph-theoretic setting, at least for "well-behaved" base graphs $G$.

# Acknowledgements

things are bound to happen whenever random quantities are close to their expectations.

However, Joel's most decisive influence on this paper is not the choice of technique, but his way of thinking. The first few times we discussed Math, I was struck by how much time he would spend pondering very specific steps in proofs, which I had mostly overlooked. I realized that this attention to detail made him understand proofs – even my own proofs – much better than I did. Trying to emulate his style made me understand at least one important thing: the main reason why Joel is so great at explaining things to others is that he first takes the time to explain them to himself.

## References

[1] Ahlswede, R. and Winter, A. (2002). Strong converse for identification via quantum channels. *IEEE Transactions on Information Theory* **48** 569–579. MR1889969

[2] Amit, A. and Linial, N. (2002). Random graph coverings, I: General theory and graph connectivity. *Combinatorica* **22** 1–18. MR1883559

[3] Amit, A. and Linial, N. (2006). Random lifts of graphs: Edge expansion. *Combinatorics, Probability and Computing* **15** 317–322. MR2216470

[4] Amit, A., Linial, N. and Matoušek, J. (2001). Random lifts of graphs: Independence and chromatic number. *Random Structures and Algorithms* **20** 1–22. MR1871947

[5] Bilu, Y. and Linial, N. (2006). Lifts, discrepancy and nearly optimal spectral gaps. *Combinatorica* **26** 495–519. MR2279667

[6] Broder, A. Z., Frieze, A. M., Suen, S. and Upfal, E. (1998). Optimal construction of edge-disjoint paths in random graphs. *SIAM Journal on Computing* **28** 541–573. MR1634360

[7] Christofides, D. and Markström, K. (2008). Expansion properties of random Cayley graphs and vertex transitive graphs via matrix martingales. *Random Structures and Algorithms* **32** 88–100. MR2371053

[8] Chung, F. R. K., Graham, R. L. and Wilson, R. M. (1989). Quasi-random graphs. *Combinatorica* **9** 345–362. MR1054011

[9] Feige, U. and Ofek, E. (2005). Spectral techniques applied to sparse random graphs. *Random Structures and Algorithms* **27** 251–275. MR2155709

[10] Friedman, J. (2003). Relative expanders or weakly relatively Ramanujan graphs. *Duke Mathematical Journal* **118** 19–35. MR1978881

[11] Friedman, J. (2008). A proof of Alon's second eigenvalue conjecture and related problems. *Memoirs of the American Mathematical Society* **195**. MR2437174

[12] Füredi, Z. and Komlós, J. (1981). The eigenvalues of random symmetric matrices. *Combinatorica* **1** 233–241. MR0637828

[13] Graham, F. C. (1997). *Spectral Graph Theory.* American Mathematical Society. MR1421568

[14] Graham, F. C. and Graham, R. L. (2008). Quasi-random graphs with given degree sequences. *Random Structures and Algorithms* **32** 1–19. MR2371048

[15] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces.* Springer-Verlag. MR1102015

[16] Linial, N. and Rozenman, E. (2005). Random lifts of graphs: Perfect matchings. *Combinatorica* **25** 407–424. MR2143247

[17] Linial, N. and Puder, D. (2010). Word maps and spectra of random graph lifts. *Random Structures and Algorithms* **37** 100–135. MR2674623

[18] Lubetsky, E., Sudakov, B. and Vu, V. (2009). Spectra of lifted Ramanujan graphs. Arxiv preprint 0911.4148.

[19] Oliveira, R. I. (2009). Concentration of the adjacency matrix and of the Laplacian in random graphs with independent edges. Arxiv preprint 0911.0600.

[20] Oliveira, R. I. (2010). Sums of random Hermitian matrices and an inequality by Rudelson. *Electronic Communications in Probability* **15** 203–212.

[21] Shamir, E. and Spencer, J. (1987). Sharp concentration of the chromatic number on random graphs $G_{n,p}$. *Combinatorica* **7** 121–129. MR0905159

[22] Tropp, J. (2010). User-friendly tail bounds for sums of random matrices. Arxiv preprint 1004.4389. MR2675131

Roberto Imbuzeiro Oliveira
IMPA
Estrada Dona Castorina, 110
Rio de Janeiro
RJ 22460-320
Brazil
*E-mail address:* rimfo@impa.br