

The spectral norm of random lifts of matrices*

Afonso S. Bandeira[†] Yunzi Ding[‡]

Abstract

We study the spectral norm of random lifts of matrices. Given an $n \times n$ symmetric matrix A , and a centered distribution π on $k \times k$ ($k \geq 2$) symmetric matrices with spectral norm at most 1, let the *matrix random lift* $A^{(k,\pi)}$ be the random symmetric $kn \times kn$ matrix $(A_{ij}X_{ij})_{1 \leq i < j \leq n}$, where X_{ij} are independent samples from π . We prove that

$$\mathbb{E}\|A^{(k,\pi)}\| \lesssim \max_i \sqrt{\sum_j A_{ij}^2} + \max_{i,j} |A_{ij}| \sqrt{\log(kn)}.$$

This result can be viewed as an extension of existing spectral bounds on random matrices with independent entries, providing further instances where the multiplicative $\sqrt{\log n}$ factor in the Non-Commutative Khintchine inequality can be removed.

As a direct application of our result, we prove an upper bound of $2(1 + \epsilon)\sqrt{\Delta} + O(\sqrt{\log(kn)})$ on the new eigenvalues for random k -lifts of a fixed $G = (V, E)$ with $|V| = n$ and maximum degree Δ , compared to the previous result of $O(\sqrt{\Delta \log(kn)})$ by Oliveira [Oli10a] and the recent breakthrough by Bordenave and Collins [BC19] which gives $2\sqrt{\Delta - 1} + o(1)$ as $k \rightarrow \infty$ for Δ -regular graph G .

Keywords: concentration inequality; random matrix theory; matrix lifts.

MSC2020 subject classifications: 60C99.

Submitted to ECP on August 22, 2020, final version accepted on June 28, 2021.

Supersedes arXiv:2006.06505.

1 Introduction

1.1 The Non-Commutative Khintchine inequality

The Non-Commutative Khintchine (NCK) inequality, originally introduced by Lust-Piquard and Pisier [Pis03], is one of the simplest tools for understanding the spectrum of matrix series, namely

$$X = \sum_{i=1}^N \gamma_i A_i \tag{1.1}$$

where A_i ($i = 1, 2, \dots, N$) are $n \times n$ real symmetric matrices and γ_i ($i = 1, 2, \dots, N$) are i.i.d. random variables, usually assumed gaussian or Rademacher. The inequality is stated as follows.

*Afonso S. Bandeira partially supported by NSF grants DMS-1712730 and DMS-1719545, and by a grant from the Sloan Foundation. Yunzi Ding partially supported by NSF grant DMS-1712730.

[†]Department of Mathematics, ETH Zurich, Switzerland. E-mail: bandeira@math.ethz.ch. Part of this work was done while ASB was with the Department of Mathematics at the Courant Institute of Mathematical Sciences, and the Center for Data Science, at New York University.

[‡]Department of Mathematics, Courant Institute of Mathematical Sciences, New York University, USA. E-mail: yding@nyu.edu.

Theorem 1.1 (Non-Commutative Khintchine (NCK) inequality). *Let A_1, A_2, \dots, A_N be $n \times n$ symmetric matrices and $\gamma_1, \gamma_2, \dots, \gamma_N$ be i.i.d. $\mathcal{N}(0, 1)$ random variables, then*

$$\mathbb{E} \left\| \sum_{i=1}^N \gamma_i A_i \right\| \leq \sigma \sqrt{2 + 2 \log(2n)}, \quad \text{where } \sigma := \left(\left\| \sum_{i=1}^N A_i^2 \right\| \right)^{1/2}. \quad (1.2)$$

The NCK inequality and other phenomena of *matrix concentration* have been proven under various settings and extensively studied in [Oli10b, Tro12, Tro15]. One particularly important application of matrix concentration is on the spectra of random matrices with independent entries. These random matrices can be represented as matrix series upon a direct entry-wise decomposition, as we show below.

1.2 Random matrices with independent entries

The study of random matrices with independent entries traces back to the seminal work by Wigner [Wig58]. For Wigner matrices (real symmetric or Hermitian random matrices with independent mean-zero and unit variance entries), a long line of work has established a comprehensive understanding towards its spectral properties over the past decades (see, for example, [FK81, BY88, AGZ10, Tao12]). One of the most important results is the Wigner semicircle law: for $n \times n$ Wigner matrix X , $\mathbb{E}\|X\|/\sqrt{n} \rightarrow 2$ and the spectrum of X converges to the semicircle $\frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{\{-2 \leq x \leq 2\}}$ as $n \rightarrow \infty$.

Random matrices with different variances on each of the independent entries, for instance real symmetric $X \in \mathbb{R}^{n \times n}$ with $X_{ij} \sim \mathcal{N}(0, b_{ij}^2)$ for $1 \leq i \leq j \leq n$, have also been studied [DS01, RV10, Ver10]. With the NCK inequality, the following estimate can be obtained:

$$\mathbb{E}\|X\| \lesssim \sigma \sqrt{\log n}, \quad \text{where } \sigma := \max_i \sqrt{\sum_j b_{ij}^2}. \quad (1.3)$$

Here $A \lesssim B$ (respectively, $A \gtrsim B$) refer to $A \leq CB$ (respectively, $A \geq CB$) for some absolute positive constant C . The definition of σ is consistent with (1.2) upon writing X as the following matrix series:

$$X = \sum_{i=1}^n \gamma_{ii} b_{ii} E_{ii} + \sum_{1 \leq i < j \leq n} \gamma_{ij} b_{ij} (E_{ij} + E_{ji}). \quad (1.4)$$

Here $E_{ij} := e_i e_j^\top$. One may immediately notice that the bound (1.3) is not sharp for Wigner matrices with i.i.d. standard gaussian entries, since it gives $\mathbb{E}\|X\| \lesssim \sqrt{n \log n}$ rather than $\mathbb{E}\|X\| \sim \sqrt{n}$. In fact, a recent improvement for matrices with independent entries [BvH16] yields

$$\mathbb{E}\|X\| \lesssim \sigma + \sigma_* \sqrt{\log n}, \quad \text{where } \sigma_* := \max_{ij} |b_{ij}|. \quad (1.5)$$

This upper bound is sharp as a matching lower bound $\mathbb{E}\|X\| \gtrsim \sigma + \sigma_* \sqrt{\log n}$ is also given in [BvH16] under mild assumptions on the b_{ij} 's. Further refinements, that hold for general b_{ij} 's, have been recently obtained [vH17, LvHY18, BGBK⁺20].

1.3 Improving the NCK inequality

The gap between the NCK bound (1.3) and its improvement (1.5) demonstrates the sub-optimality of the NCK inequality in some settings. In fact, many improvements to bounds obtained via the NCK inequality are known under various settings. Seginer [Seg00] improved the NCK bound for random matrices with independent uniformly bounded entries. Exploiting the non-commutativity among the A_i 's, Tropp [Tro18]

proved the following upper bound for the series (1.1), which improved the multiplicative factor on σ from $\sqrt{\log n}$ to $\sqrt[4]{\log n}$:

$$\mathbb{E}\|X\| \lesssim \sigma \sqrt[4]{\log n} + \omega \sqrt{\log n}$$

where the alignment parameter ω is defined as

$$\omega := \max_{Q_1, Q_2, Q_3 \in U_n} \left\| \sum_{i,j=1}^n A_i Q_1 A_j Q_2 A_i Q_3 A_j \right\|^{1/4},$$

here U_n denotes the group of $n \times n$ unitary matrices (the paper [Tro18] considers Hermitian A_i 's, while in this paper we focus on the real symmetric setting). In fact, a bound which replaces the multiplicative $\sqrt{\log n}$ factor in the NCK inequality by an additive factor has been hypothesized in many different forms [Tro12, Ban15, BvH16, vH17, LvHY18, Tro18]. For the matrix series (1.1), define the “weak variance” as

$$\sigma_* = \left(\max_{\|v\|=1} \sum_{i=1}^N (v^\top A_i v)^2 \right)^{1/2}, \tag{1.6}$$

a possible improvement to Theorem 1.2 could be

$$\mathbb{E} \left\| \sum_{i=1}^N \gamma_i A_i \right\| \leq C(\sigma + \sigma_* \sqrt{\log n}). \tag{1.7}$$

Note that $\sigma_* \leq \sigma$ by a simple application of the Cauchy-Schwarz inequality. For random matrices with independent $X_{ij} \sim \mathcal{N}(0, b_{ij}^2)$ for $1 \leq i \leq j \leq n$, upon writing it as a matrix series as in (1.4), we have

$$X = \sum_{i=1}^N \gamma_i A_i, \quad \text{where } \{A_i\}_{i=1}^N = \{b_{ii} E_{ii}\}_{i=1}^n \cup \{b_{ij}(E_{ij} + E_{ji})\}_{1 \leq i < j \leq n}.$$

It is not hard to show that

$$\max_{\|v\|=1} \left(\sum_{i=1}^N (v^\top A_i v)^2 \right)^{1/2} \asymp \max_{ij} |b_{ij}|,$$

thus σ_* defined in (1.6) is consistent with the quantity defined in (1.5) in that they differ only by a multiplicative constant, and the proposed improvement (1.7) is indeed true in the case of random matrices with independent entries due to [BvH16].

In this paper, we show another class of examples, in which we improve the bound given by the NCK inequality (1.2) by replacing the multiplicative log factor with an additive factor, as is the case in the conjectured bound (1.7). In the following context, all the matrices we consider are real. As an extension to random matrices with independent entries, we consider the operation of *matrix lifts*, in which each entry of an underlying deterministic matrix is replaced by the product of itself and a random $k \times k$ matrix, as described in the following definition.

Definition 1.2 (Matrix lifts). *Let A be an $n \times n$ symmetric matrix with zero diagonal entries, and π be a measure supported on $k \times k$ matrices. Define the (k, π) lift of A , denoted $A^{(k, \pi)}$, as follows:*

- Draw i.i.d. samples $\{\Pi_{ij}\}_{1 \leq i < j \leq n}$ from π ; for $1 \leq i < j \leq n$, denote $\Pi_{ji} := \Pi_{ij}^\top$;
- For $1 \leq i \leq n$, denote $\Pi_{ii} := 0_k$;

- For all $1 \leq i, j \leq n$, replace A_{ij} with the matrix $A_{ij}\Pi_{ij}$.

The resulting matrix is a $kn \times kn$ symmetric random matrix, which can be written as

$$A^{(k,\pi)} = \sum_{1 \leq i, j \leq n} A_{ij}(E_{ij} \otimes \Pi_{ij}) \tag{1.8}$$

where the symbol “ \otimes ” on the RHS denotes the Kronecker product of matrices.

The main theorem of this paper is the following bound:

Theorem 1.3. *Let A be a symmetric $n \times n$ matrix ($n \geq 2$) with zero diagonal entries. Suppose π is a centered measure supported on $k \times k$ matrices with spectral norm at most 1. Then there exists a universal constant C , such that for any $\epsilon \in (0, 1/2]$,*

$$\mathbb{E}\|A^{(k,\pi)}\| \leq 2(1 + \epsilon)\sigma + \frac{C}{\sqrt{\log(1 + \epsilon)}}\sigma_*\sqrt{\log(kn)}. \tag{1.9}$$

where

$$\sigma := \max_i \sqrt{\sum_j A_{ij}^2}, \quad \sigma_* := \max_{ij} |A_{ij}|.$$

Note that Definition 1.2 and Theorem 1.3 only apply to base matrices A with zero diagonal entries. A base matrix with possibly non-zero diagonal entries can be handled by splitting it in its diagonal and non-diagonal parts and using triangular inequality in the result random matrices.

Remark 1.4. Upon taking $A = \{b_{ij}\}, k = 1$ and $\pi = \text{Uniform}\{\pm 1\}$, Definition 1.2 and Theorem 1.3 include as a special case the real symmetric random matrix $X \in \mathbb{R}^{n \times n}$ with $X_{ij} = \epsilon_{ij}b_{ij}$, where $\{b_{ij}\}$ are given and ϵ_{ij} are independent Rademacher random variables for $1 \leq i \leq j \leq n$, that is, $\mathbb{P}[\epsilon_{ij} = \pm 1] = 1/2$. Since [BvH16] showed that the bound $O(\sigma + \sigma_*\sqrt{\log n})$ captures the optimal scaling of $\mathbb{E}\|X\|$ with respect to σ and $\sigma_*\sqrt{\log n}$ and is in general unimprovable, this implies the same for our bound (1.9) on $\mathbb{E}\|A^{(k,\pi)}\|$. However, Theorem 1.3 does not directly imply the bound (1.5), since gaussian random variables are not compactly supported.

Besides the $k = 1$ case, Theorem 1.3 is also interesting with natural choices such as π being the Haar measure on the orthogonal group $O(k)$ or special orthogonal group $SO(k)$. One particular application is an estimate on the spectrum of random lifts of graphs, which we discuss below.

1.4 Application: random lifts of graphs

Given an undirected graph $G = (V, E)$ and an integer $k \geq 2$, the random k -lift of G , denoted $G^{(k)}$, is obtained by replacing each vertex $v \in V$ by k new vertices, and each edge $e = (v_1, v_2)$ by a random $k \times k$ bipartite matching between the k new vertices corresponding to v_1 and those corresponding to v_2 . Here “random” refers to a uniform choice on all $k!$ possible bipartite matchings. We denote A and $A^{(k)}$ the adjacency matrix of G and $G^{(k)}$, respectively.

Previous studies on the k -lifts of graphs, under the setting of fixed G and $k \rightarrow \infty$, have revealed many properties of the resulting random graph, such as connectivity [AL02], chromatic number [ALM02], edge expansion [AL06] and the existence of perfect matching [LR05]. The spectrum of random k -lifts, namely the new spectrum introduced in the lifting process

$$\max_{\eta \in \text{spec}(A^{(k)}) \setminus \text{spec}(A)} |\eta| = \|A^{(k)} - \mathbb{E}A^{(k)}\| \tag{1.10}$$

was studied by Friedman [Fri03] via the trace method, who showed that with a random d -regular graph as the base graph, as $k \rightarrow \infty$, (1.10) is $O(d^{3/4})$ with high probability.

He also conjectured the tight bound $2\sqrt{d-1} + o(1)$. The high probability upper bound on (1.10) was improved by Linial and Puder to $O(d^{2/3})$ in [LP10], then by Lubetzky, Sudakov and Vu [LSV11] to $O(\sqrt{d} \log d)$ in the case that the second eigenvalue of the base graph is $O(\sqrt{d})$. Later, Addario-Berry and Griffiths [ABG10] and Puder [Pud15] proved that (1.10) is $O(\sqrt{d})$, the latter giving $2\sqrt{d-1} + O(1)$ as an upper bound. Since then, various extensions or alternative proofs of the bound on (1.10) of the scale $O(\sqrt{d})$ (some under slightly different settings) have been carried out with different combinatorial and probabilistic techniques, for example, in [FK14, BLM15, Bor19, ACKM17].

We should notice that the above line of work adopted the asymptotic regime $k \rightarrow \infty$ and the setting that the base graph is taken randomly over all d -regular graphs on n vertices. In fact, in the case that the base graph is a fixed d -regular graph and $k \in \mathbb{N}_+$ is fixed, (1.10) is not always upper bounded by $O(\sqrt{d})$. As a counterexample (see [BvH16], Remark 4.8): consider G the union of n/s cliques of s vertices each, with no edges between different cliques; here $s = \lceil \sqrt{\log n} \rceil$, and we assume that n/s is an integer for simplicity. Seginer [Seg00] showed that

$$\mathbb{E}\|A^{(2)} - \mathbb{E}A^{(2)}\| \sim \sqrt{\log n},$$

whereas the $O(\sqrt{d})$ bound would incorrectly predict that LHS is $O(\log^{1/4}(n))$.

Another line of work considers a fixed base graph G with maximum degree Δ without assuming its randomness. Making use of matrix concentration, Oliveira [Oli10a] obtained a high probability upper bound of $O(\sqrt{\Delta \log(kn)})$ on (1.10). The most recent advancement by Bordenave and Collins [BC19] considered the k -lifts problem under a much more general framework, and proved that (1.10) is $2\sqrt{d-1} + o(1)$ for any d -regular base graph G as $k \rightarrow \infty$, finally settling Friedman’s conjecture even without assuming the randomness of the base graph.

In what follows, we manage to improve the bound in [Oli10a] by removing the multiplicative factor $\sqrt{\log(kn)}$, replacing it with an additive factor. We also improve the constant factor before $\sqrt{\Delta}$ down to 2, in consistence with Friedman’s theorem and [BC19]. In the large k limit, our bound is weaker than [BC19] by an additive $\sqrt{\log(kn)}$ factor. For $k = 2$, an additive $\sqrt{\log n}$ factor is needed as illustrated by a counterexample due to Seginer [Seg00]. However, our result does not capture the correct dependence on k , namely concentration arising from large k . Note that we are only using a slight modification of the moment method, compared to the sophisticated combinatorial technique in [BC19].

Theorem 1.5. *Let A be the adjacency matrix of $G = (V, E)$ with $|V| = n$ and $\max \deg(G) = \Delta$, and $A^{(k)}$ be the corresponding random k -lift. Then there exists a universal constant C , such that for any $\epsilon \in (0, 1/2]$*

$$\mathbb{E}\|A^{(k)} - \mathbb{E}A^{(k)}\| \leq 2(1 + \epsilon)\sqrt{\Delta} + \frac{C}{\sqrt{\log(1 + \epsilon)}}\sqrt{\log(kn)}. \tag{1.11}$$

Our bound is essentially $2(1 + \epsilon)\sqrt{\Delta}$ as long as $\Delta \gg \log(kn)$, i.e. the base graph G is not too sparse. The proof of Theorem 1.5 will follow from our main result, Theorem 1.3.

Notation

In this paper, for positive quantities A and B , $A \lesssim B$ and $A \gtrsim B$ respectively refer to $A \leq CB$ and $A \geq CB$ for some absolute positive constant C . For $x \in \mathbb{R}$, $\lceil x \rceil$ denotes the minimum integer that is larger than or equal to x .

2 Proof of main results

In this section, we carry out the proof of Theorems 1.3 and 1.5. We begin with the following comparison argument which links $A^{(k,\pi)}$ to an auxiliary Wigner matrix. This argument is a modification of Proposition 2.1 in [BvH16], and the auxiliary matrix Y_r is same as in the proof of Theorem 4.8 in [LvHY18].

Proposition 2.1. *Let Y_r be the $r \times r$ symmetric matrix with zero diagonal and*

$$Y_{ij} = \begin{cases} \sqrt{3}, & w.p. \frac{1}{4} \\ -\frac{1}{\sqrt{3}}, & w.p. \frac{3}{4} \end{cases}$$

independently for all $1 \leq i < j \leq r$. Under the setting of Theorem 1.3, suppose $\sigma_ \leq 1$, then for every $p \in \mathbb{N}_+$ there holds*

$$\mathbb{E} \text{Tr} \left[(A^{(k,\pi)})^{2p} \right] \leq \frac{kn}{\lceil \sigma^2 \rceil + p} \mathbb{E} \text{Tr} \left[Y_{\lceil \sigma^2 \rceil + p}^{2p} \right].$$

To carry out the proof of Proposition 2.1, we start with a set of standard notations adopted from [FK81] and [BvH16]. Following the representation (1.8), a direct expansion of $(A^{(k,\pi)})^{2p}$ yields

$$\mathbb{E} \text{Tr} \left[(A^{(k,\pi)})^{2p} \right] = \sum_{u_1, u_2, \dots, u_{2p} \in [n]} \left(\prod_{j=1}^{2p} A_{u_j u_{j+1}} \right) \mathbb{E} \text{Tr} \left[\prod_{j=1}^{2p} \Pi_{u_j u_{j+1}} \right]. \quad (2.1)$$

Let $G_n = ([n], E_n)$ be the complete graph on n points. A cycle $u_1 \rightarrow u_2 \rightarrow \dots \rightarrow u_{2p} \rightarrow u_1$ of length $2p$, where $u_i \in [n]$ for all $1 \leq i \leq 2p$ ($u_{2p+1} := u_1$), is identified as $\mathbf{u} = (u_1, \dots, u_{2p}) \in [n]^{2p}$. Since $\mathbb{E} [\Pi_{ij}] = 0$ for any $1 \leq i, j \leq n$, in the sum of (2.1) we only need to consider cycles with each edge appearing at least twice.

We call the *shape* of a cycle \mathbf{u} , denoted $\mathbf{s}(\mathbf{u})$, a relabeling of the vertices in the order of their appearance. For example, the shape of $\mathbf{u} = (4, 7, 2, 7, 9, 4, 5, 4)$ is $(1, 2, 3, 2, 4, 1, 5, 1)$. The following set is a collection of all cycles of shapes that contribute to the sum in (2.1):

$$\mathcal{S}_{2p} := \{ \mathbf{s}(\mathbf{u}) : \mathbf{u} \text{ is a cycle of length } 2p \text{ with each edge appearing at least twice} \}.$$

For the sake of convenience, we also define the set of cycles with fixed shape and starting point as

$$\Gamma_{\mathbf{s}, u} := \{ \mathbf{u} \in [n]^{2p} : \mathbf{s}(\mathbf{u}) = \mathbf{s}, u_1 = u \}.$$

The *span* of a shape \mathbf{s} , denoted by $m(\mathbf{s})$, is the largest index in its representation, also the number of distinct vertices any cycle of shape \mathbf{s} visits. A direct observation is $m(\mathbf{s}) \leq p + 1$ for any $\mathbf{s} \in \mathcal{S}_{2p}$.

Proof of Proposition 2.1. Following the expansion (2.1) we have

$$\begin{aligned} \mathbb{E} \text{Tr} \left[(A^{(k,\pi)})^{2p} \right] &= \sum_{\mathbf{s} \in \mathcal{S}_{2p}} \sum_{u \in [n]} \sum_{\mathbf{u} \in \Gamma_{\mathbf{s}, u}} \left(\prod_{j=1}^{2p} A_{u_j u_{j+1}} \right) \mathbb{E} \text{Tr} \left[\prod_{j=1}^{2p} \Pi_{u_j u_{j+1}} \right] \\ &\leq k \sum_{\mathbf{s} \in \mathcal{S}_{2p}} \sum_{u \in [n]} \sum_{\mathbf{u} \in \Gamma_{\mathbf{s}, u}} \left(\prod_{j=1}^{2p} |A_{u_j u_{j+1}}| \right) \\ &\leq kn \sum_{\mathbf{s} \in \mathcal{S}_{2p}} \sigma^{2(m(\mathbf{s})-1)} \end{aligned} \quad (2.2)$$

where the first inequality follows from $\|\prod_{j=1}^{2p} \Pi_{u_j u_{j+1}}\| \leq 1$ and therefore $\text{Tr} \left[\prod_{j=1}^{2p} \Pi_{u_j u_{j+1}} \right] \leq k$; and the second inequality owes to the fact that, under $\sigma_* \leq 1$, for any $u \in [n]$ and $\mathbf{s} \in \mathcal{S}_{2p}$, Lemma 2.5 in [BvH16] gives

$$\sum_{\mathbf{u} \in \Gamma_{\mathbf{s}, u}} \left(\prod_{j=1}^{2p} |A_{u_j u_{j+1}}| \right) \leq \sigma^{2(m(\mathbf{s})-1)}.$$

Meanwhile, for any positive integer $r > p$, for the auxiliary random matrix Y_r we have

$$\begin{aligned} \mathbb{E} \text{Tr} [Y_r^{2p}] &= \sum_{\mathbf{u} \in [r]^{2p}} \mathbb{E} \text{Tr} \left[\prod_{j=1}^{2p} Y_{u_j u_{j+1}} \right] \\ &\geq \sum_{\mathbf{s} \in \mathcal{S}_{2p}} \sum_{u \in [r]} \sum_{\mathbf{u} \in \Gamma_{\mathbf{s}, u}} \mathbb{E} \text{Tr} \left[\prod_{j=1}^{2p} Y_{u_j u_{j+1}} \right] \\ &= \sum_{\mathbf{s} \in \mathcal{S}_{2p}} |\{\mathbf{u} \in [r]^{2p} : \mathbf{s}(\mathbf{u}) = \mathbf{s}\}| \cdot \mathbb{E} \text{Tr} \left[\prod_{j=1}^{2p} Y_{u_j u_{j+1}} \right] \\ &\geq \sum_{\mathbf{s} \in \mathcal{S}_{2p}} r(r-1) \cdots (r-m(\mathbf{s})+1) \cdot 1 \end{aligned}$$

The last inequality follows from the observation that $\mathbb{E}[Y_{ij}^m] \geq 1$ for all $m \geq 2$. Now choosing $r = \lceil \sigma^2 \rceil + p$, noting that $m(\mathbf{s}) \leq p + 1$ for all $\mathbf{s} \in \mathcal{S}_{2p}$, we have

$$\mathbb{E} \text{Tr} [Y_{\lceil \sigma^2 \rceil + p}^{2p}] \geq (\lceil \sigma^2 \rceil + p) \sum_{\mathbf{s} \in \mathcal{S}_{2p}} \sigma^{2(m(\mathbf{s})-1)}. \tag{2.3}$$

Comparing (2.2) with (2.3) yields the result. □

The following lemma gives an upper bound on the moments of the auxiliary random matrix Y_r .

Lemma 2.2. *For Y_r defined in Proposition 2.1, there exists an absolute constant C , such that for any positive integer $p \geq 2$, there holds*

$$\mathbb{E} \text{Tr} [Y_r^{2p}] \leq r(2\sqrt{r} + C\sqrt{p})^{2p}. \tag{2.4}$$

Proof. Since

$$\mathbb{E} \text{Tr} [Y_r^{2p}] \leq r \mathbb{E} [\|Y_r\|^{2p}],$$

we only need to show that there exists an absolute constant C , such that for $p \geq 2$,

$$\mathbb{E} [\|Y_r\|^{2p}] \leq (2\sqrt{r} + C\sqrt{p})^{2p}. \tag{2.5}$$

The proof of (2.5) is contained in the proof of Theorem 4.8 in [LvHY18], so we do not repeat it here. The main steps of the proof are a norm bound for Wigner matrices with non-symmetrically distributed entries followed by Talagrand’s concentration inequality. □

Proof of Theorem 1.3. By Proposition 2.1 and Lemma 2.2, assuming $\sigma_* \leq 1$, we know that for any positive integer $p \geq 2$,

$$\begin{aligned} \mathbb{E} \|A^{(k, \pi)}\| &\leq \left(\mathbb{E} \text{Tr} [(A^{(k, \pi)})^{2p}] \right)^{1/2p} \\ &\leq \left(\frac{kn}{\lceil \sigma^2 \rceil + p} \mathbb{E} \text{Tr} [Y_{\lceil \sigma^2 \rceil + p}^{2p}] \right)^{1/2p} \\ &\leq (kn)^{1/2p} \left(2\sqrt{\lceil \sigma^2 \rceil + p} + C\sqrt{p} \right) \end{aligned}$$

If $kn \geq 3$, for $\alpha \geq 1$ choosing $p = \lceil \alpha \log(kn) \rceil \geq 2$ yields

$$\begin{aligned} \mathbb{E}\|A^{(k,\pi)}\| &\leq e^{1/2\alpha} \left(2\sqrt{\lceil \sigma^2 \rceil + \lceil \alpha \log(kn) \rceil} + C\sqrt{\lceil \alpha \log(kn) \rceil} \right) \\ &\leq e^{1/2\alpha} \left(2\sigma + 2\sqrt{\alpha \log(kn) + 2} + C\sqrt{\alpha \log(kn) + 1} \right). \end{aligned}$$

Denote $e^{1/2\alpha} = 1 + \epsilon$. Since $n \geq 2$, $k \geq 1$, and $\epsilon \leq 1/2$ implies $\alpha \geq 1$, we have $2 < 3 \log 2 \leq 3\alpha \log(kn)$, so

$$\mathbb{E}\|A^{(k,\pi)}\| \leq 2(1 + \epsilon)\sigma + (1 + \epsilon)(4 + 2C)\sqrt{\frac{\log(kn)}{2 \log(1 + \epsilon)}}.$$

The remaining case $kn < 3$ can only happen when $n = 2$ and $k = 1$. In this case π is supported on $[-1, 1]$, and we can directly estimate

$$\mathbb{E}\|A^{(k,\pi)}\| \leq \mathbb{E}\|A^{(k,\pi)}\|_F \leq 2 < 2\sqrt{2 \log(kn)}. \quad \square$$

The spectral bound of random k -lifts of graphs follows as an immediate corollary.

Proof of Theorem 1.5. Denote $\text{Perm}(k)$ the collection of all $k \times k$ permutation matrices, and

$$G_k := \left\{ \Pi - \frac{1}{k} J_k : \Pi \in \text{Perm}(k) \right\}$$

where J_k is the $k \times k$ matrix with all entries 1. It is easy to verify that $\|X\| \leq 1$ for any $X \in G_k$. Moreover, the adjacency matrix A has $\sigma^2 \leq \Delta$ and $\sigma_* \leq 1$ by definition. Thus it follows from Theorem 1.3 that

$$\begin{aligned} \mathbb{E}\|A^{(k)} - \mathbb{E}A^{(k)}\| &= \mathbb{E}\|A^{(k, \text{Unif}(G_k))}\| \\ &\leq 2(1 + \epsilon)\sqrt{\Delta} + \frac{C}{\sqrt{\log(1 + \epsilon)}}\sqrt{\log(kn)}. \end{aligned} \quad \square$$

Remark 2.3. In the above proof, we applied Theorem 1.3 on $\pi = \text{Unif}(G_k)$, where G_k is the centered version of $\text{Perm}(k)$. One may expect that, under the setting of Theorem 1.3 without assuming π is centered, there still holds

$$\mathbb{E}\|A^{(k,\pi)} - \mathbb{E}A^{(k,\pi)}\| \leq C \left(\sigma + \sigma_* \sqrt{\log(kn)} \right). \quad (2.6)$$

Though we do not have a counterexample for (2.6), we must point out that (2.6) only follows from Theorem 1.3 when $\|X - \mathbb{E}_\pi X\| \leq 1$ for every $X \in \text{supp}(\pi)$.

We note that the proof of Theorem 1.3 is not exploiting any potential structure of the “lifting matrices” Π_{ij} . In fact, this may explain why Theorem 1.5 is worse than the result in [BC19] by an additive $\sqrt{\log(kn)}$ factor in the large k limit for d -regular base graphs. One may be able to obtain a stronger result, for instance $\mathbb{E}\|A^{(k,\pi)} - \mathbb{E}A^{(k,\pi)}\| \leq 2\sqrt{\Delta} + o_k(1)$, with a more careful analysis considering that Π_{ij} are permutation matrices.

References

- [ABG10] Louigi Addario-Berry and Simon Griffiths. The spectrum of random lifts. *arXiv preprint arXiv:1012.4097*, 2010.
- [ACKM17] Naman Agarwal, Karthekeyan Chandrasekaran, Alexandra Kolla, and Vivek Madan. On the expansion of group-based lifts. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

- [AGZ10] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118. Cambridge university press, 2010.
- [AL02] Alon Amit and Nathan Linial. Random graph coverings i: General theory and graph connectivity. *Combinatorica*, 22(1):1–18, 2002.
- [AL06] Alon Amit and Nathan Linial. Random lifts of graphs: edge expansion. *Combinatorics, Probability and Computing*, 15(3):317–332, 2006.
- [ALM02] Alon Amit, Nathan Linial, and Jiří Matoušek. Random lifts of graphs: independence and chromatic number. *Random Structures & Algorithms*, 20(1):1–22, 2002.
- [Ban15] Afonso S Bandeira. Ten lectures and forty-two open problems in the mathematics of data science. *Lecture Notes*, 2015.
- [BC19] Charles Bordenave and Benoît Collins. Eigenvalues of random lifts and polynomials of random permutation matrices. *Annals of Mathematics*, 190(3):811–875, 2019.
- [BGBK⁺20] Florent Benaych-Georges, Charles Bordenave, Antti Knowles, et al. Spectral radii of sparse random matrices. In *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, volume 56, pages 2141–2161. Institut Henri Poincaré, 2020.
- [BLM15] Charles Bordenave, Marc Lelarge, and Laurent Massoulié. Non-backtracking spectrum of random graphs: community detection and non-regular ramanujan graphs. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1347–1357. IEEE, 2015.
- [Bor19] Charles Bordenave. A new proof of friedman’s second eigenvalue theorem and its extension to random lifts. In *Annales scientifiques de l’Ecole normale supérieure*, 2019.
- [BvH16] Afonso S Bandeira and Ramon van Handel. Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *The Annals of Probability*, 44(4):2479–2506, 2016.
- [BY88] Zhi-Dong Bai and Yong-Qua Yin. Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a wigner matrix. *The Annals of Probability*, pages 1729–1741, 1988.
- [DS01] Kenneth R Davidson and Stanislaw J Szarek. Local operator theory, random matrices and banach spaces. *Handbook of the geometry of Banach spaces*, 1(317-366):131, 2001.
- [FK81] Zoltán Füredi and János Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.
- [FK14] Joel Friedman and David-Emmanuel Kohler. The relativized second eigenvalue conjecture of alon. *arXiv preprint arXiv:1403.3462*, 2014.
- [Fri03] Joel Friedman. Relative expanders or weakly relatively ramanujan graphs. *Duke Mathematical Journal*, 118(1):19–35, 2003.
- [LP10] Nati Linial and Doron Puder. Word maps and spectra of random graph lifts. *Random Structures & Algorithms*, 37(1):100–135, 2010.
- [LR05] Nathan Linial and Eyal Rozenman. Random lifts of graphs: perfect matchings. *Combinatorica*, 25(4):407–424, 2005.
- [LSV11] Eyal Lubetzky, Benny Sudakov, and Van Vu. Spectra of lifted ramanujan graphs. *Advances in Mathematics*, 227(4):1612–1645, 2011.
- [LvHY18] Rafał Łatała, Ramon van Handel, and Pierre Youssef. The dimension-free structure of nonhomogeneous random matrices. *Inventiones mathematicae*, 214(3):1031–1080, 2018.
- [Oli10a] Roberto Imbuzeiro Oliveira. The spectrum of random k -lifts of large graphs (with possibly large k). *Journal of Combinatorics*, 1(3):285–306, 2010.
- [Oli10b] Roberto Imbuzeiro Oliveira. Sums of random hermitian matrices and an inequality by rudelson. *Electronic Communications in Probability*, 15:203–212, 2010.
- [Pis03] Gilles Pisier. *Introduction to operator space theory*, volume 294. Cambridge University Press, 2003.

- [Pud15] Doron Puder. Expansion of random graphs: New proofs, new results. *Inventiones mathematicae*, 201(3):845–908, 2015.
- [RV10] Mark Rudelson and Roman Vershynin. Non-asymptotic theory of random matrices: extreme singular values. In *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) (In 4 Volumes) Vol. I: Plenary Lectures and Ceremonies Vols. II–IV: Invited Lectures*, pages 1576–1602. World Scientific, 2010.
- [Seg00] Yoav Seginer. The expected norm of random matrices. *Combinatorics, Probability and Computing*, 9(2):149–166, 2000.
- [Tao12] Terence Tao. *Topics in random matrix theory*, volume 132. American Mathematical Soc., 2012.
- [Tro12] Joel A Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of computational mathematics*, 12(4):389–434, 2012.
- [Tro15] Joel A Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends in Machine Learning*, 8(1-2):1–230, 2015.
- [Tro18] Joel A Tropp. Second-order matrix concentration inequalities. *Applied and Computational Harmonic Analysis*, 44(3):700–736, 2018.
- [Ver10] Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027*, 2010.
- [vH17] Ramon van Handel. On the spectral norm of gaussian random matrices. *Transactions of the American Mathematical Society*, 369(11):8161–8178, 2017.
- [Wig58] Eugene P Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, pages 325–327, 1958.

Acknowledgments. We are grateful to Ramon van Handel for comments on an early version of the paper, in particular for pointing us to the latest results on graph k -lifts in [BC19], for directing us to the proof of Theorem 4.8 in [LvHY18] which allowed us to improve the constant factor before σ in Theorem 1.3 to 2, and for making us aware of recent efforts to improve the NCK inequality under more general settings. We would also like to thank Jiedong Jiang, Eyal Lubetzky, Ruedi Suter and Joel Tropp for helpful discussions.