BOUNDING NORMS OF LOCALLY RANDOM MATRICES

Dhruv Medarametla

Abstract. Recently, several papers proving lower bounds for the performance of the Sum Of Squares Hierarchy on the planted clique problem have been published. A crucial part of all four papers is probabilistically bounding the norms of certain "locally random" matrices. In these matrices, the entries are not completely independent of each other, but rather depend upon a few edges of the input graph. In this paper, we study the norms of these locally random matrices.

We start by bounding the norms of simple locally random matrices, whose entries depend on a bipartite graph H and a random graph G; we then generalize this result by bounding the norms of complex locally random matrices, matrices based off of a much more general graph H and a random graph G. For both cases, we prove almost-tight probabilistic bounds on the asymptotic behavior of the norms of these matrices.

1. Introduction

Recently, several papers proving lower bounds for the performance of the Sum Of Squares Hierarchy on the planted clique problem have been published [12,17,22,27]. These papers utilize many different techniques in order to prove these results; however, a crucial part of all four papers is probabilistically bounding the norms of certain "locally random" matrices. In these matrices, the entries are not completely independent of each other, but rather depend upon a few edges of the input graph. In this paper, we study the norms of these locally random matrices. To the best of our knowledge, these matrices have not previously been studied in random matrix theory.

1.1. **Background and Motivation.**

An important problem in the field of algorithm design is the *planted clique problem*, introduced by Jerrum [18] and Kucera [20]. Let $G(n,p)$ denote the Erdös-Rényi model for random graphs on n labeled vertices, where each edge between any two vertices of the graph is included with probability p. Then, in the planted clique problem, we are given a graph $G = G(n, 1/2, k)$, created by first choosing a random graph from $G(n, 1/2)$ and then placing a clique of size k randomly in the graph for $k \gg \log(n)$. The goal of the problem is then to recover the clique for as small a k as possible given G. Though the problem was proposed over 20 years ago, the current best polynomial time algorithms can only solve the planted clique problem for $k = \Theta(\sqrt{n})$ [2]. Though a near-linear algorithm was developed for this value of k in [11], the issue of solving the problem for smaller values of k has received significant attention.

The planted clique problem has also received attention due to its applications in several other fields of mathematics. It has been used to explain the difficulty of sparse principal component detection in [10], and its level of difficulty has led to proposals to base cyrpto systems on variants of it in [3]. In addition, the planted clique problem has been utilized to show that evaluation of certain financial derivatives is hard in [4], and it has also been used to show that testing k-wise independence is hard near the information theoretic limit in [1]. The planted clique problem also has applications outside of just mathematics; it is closely related to the problem of finding large communities of people in social networks, and is studied in molecular biology with respect to signal finding, a process dedicated to finding patterns in DNA sequences, in [26]. Due to the versatility and the significance of the planted clique problem, understanding its difficulty with respect to algorithmic design is an important open problem.

A potential technique for solving the planted clique problem is the *Sum Of Squares Hierarchy*. The Sum Of Squares Hierarchy, or the SOS Hierarchy, is a generalization of semi-definite programming, a common and powerful tool for obtaining approximate solutions to combinatorial optimization problems. As with other hierarchies, the SOS hierarchy works by presenting progressively stronger convex relaxations for combinatorial optimization problems parametrized by the number of rounds r. The SOS hierarchy is known to be very powerful. In particular, it is known that the Hierarchy captures the Goemmans-Williamson algorithm for max-cut [13] and the Arora, Rao, Vazirani algorithm for sparsest cut [6], and [9] and [16] have shown that the SOS hierarchy captures the sub-exponential algorithm for unique games of [5]. Despite this knowledge, we do not have a good understanding of the performance of the SOS Hierarchy for most problems. Most known lower bounds have their origins in [14, 15], though some of these bounds were later independently rediscovered in [28]. Bettering the understanding the power and the limitations of the SOS hierarchy is a major area of study in complexity theory, as doing so would give great insight into the powers and limitations of general computation.

1.2. Outline.

Given a bipartite graph H with partite sets of size u and v and a random graph G, one can create an $\frac{n!}{(n-u)!} \times \frac{n!}{(n-v)!}$ locally random matrix R whose entries are solely dependent upon the presence of particular edges in G (see Definition 2 for an exact definition). We first bound the norm of this matrix when our bipartite graph H is just K_2, or the complete graph on two vertices; while this case has been extensively studied, such as in [7], we investigate it in order to gain intuition for the following sections. We then move on to the general case of any bipartite graph H. We bound the norm of R in terms of u, v, q, and n, where q is the size of a minimum vertex cover of H. Finally, we study an even broader case, in which the matrix whose norm we attempt to bound is based on a much more general input graph H. We utilize the same techniques as before–namely, the trace method and certain theorems from graph theory–to calculate an upper bound for this norm. We conclude by analyzing some of the ways in which the work in this paper can be applied to prove specific cases in already-existing literature, such as in [17, 22, 27], as well as suggesting future directions.

2. Definitions and Preliminaries

This paper utilizes several terms specific to matrix theory and graph theory, most of which follow their standard definitions. We recall the following standard linear algebra definitions. Given a matrix M, we refer to the element in the i^{th} row and the j^{th} column of M by the term $M(i,j)$. In addition, given an $m \times n$ matrix M, the spectral norm of M induced by the Euclidean 2 norm is defined to be $||M|| = \max_{||v||=1} ||Mv||$, where v is an n-dimensional vector. We also define the trace of an $n \times n$ matrix M to be equal to $\operatorname{tr}(M) = \sum_{i=1}^{n} M(i,i)$. Finally, we define the eigenvalues of an $n \times n$ matrix M to be the n values λ_i such that there exists an n-dimensional vector v with $Mv = \lambda_i v$.

Because we are studying the norm of certain matrices, it is logical to develop a method to analyze these norms. Throughout this paper, we utilize what is known as the trace method to bound these norms.

Proposition 2.1. *Given a matrix A, $\sqrt[2k]{\operatorname{tr}((AA^T)^k)} \geq ||A||$.*

Proof. Note that if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the n eigenvalues of a matrix M, then $\operatorname{tr}(M^k) = \lambda_1^k + \lambda_2^k + \cdots + \lambda_n^k$. However, if $\lambda_{\max}(M)$ is the largest eigenvalue of a matrix M, then $||A|| = \sqrt{\lambda_{\max}(AA^T)}$. Therefore,

$$\sqrt[2k]{\operatorname{tr}((AA^T)^k)} \geq \sqrt[2k]{\lambda_{\max}(AA^T)^k} \geq ||A||.$$

$\square$

3. Bounding the Norm of a Singular Locally Random Matrix

Consider a random graph $G = G(n, \frac{1}{2})$ with vertices labeled from 1 to n. We then define a singular locally random matrix as follows:

Definition 1. Given a labeled random graph $(V(G), E(G)) = G = G(n, \frac{1}{2})$ with $|V| = n$, create an $n \times n$ matrix R, known as the singular locally random matrix, with the following entries:

$$R(i,j) = \begin{cases} 0 & i = j \\ 1 & (i,j) \in E(G) \\ -1 & (i,j) \notin E(G). \end{cases}$$

Note that the singular locally random matrix is closely related to the adjacency matrix of a graph; in fact, it is the additive inverse of the Seidel adjacency matrix.

Note that for all $1 \leq i, j \leq n$, $\mathbb{E}[R(i,j)] = 0$, as any edge (i,j) has probability $\frac{1}{2}$ of being included in G. In addition, note that R is symmetric, so $R = R^T$.

This type of matrix and its norm have been studied extensively already. In particular, Wigner showed in 1958 in [30] that the norm of such a matrix was $O(\sqrt{n})$; in fact, this norm was shown to be very close to $2\sqrt{n}$ with high probability. Nonetheless, we still consider this example in our report as it helps us gain intuition and better approach the following problems.

In order to find a probabilistic bound for $||R||$, we bound $\mathbb{E}[\sqrt[2k]{\mathrm{tr}((RR^T)^k)}]$. Notice that

$$\mathrm{tr}((RR^T)^k) = \mathrm{tr}(R^{2k}) = \sum_{i_1,i_2,\ldots,i_{2k}\in[1,n]} \left(\prod_{j=1}^{2k} R(i_j,i_{j+1})\right)$$

where $i_{2k+1} = i_1$ and $[1,n] = \{1,2,\ldots,n\}$. This result is a simple application of matrix multiplication. Therefore,

$$\mathbb{E}[\mathrm{tr}((RR^T)^k)] = \mathbb{E}[\mathrm{tr}(R^{2k})] = \mathbb{E}\left[\sum_{i_1,i_2,\ldots,i_{2k}\in[1,n]} \left(\prod_{j=1}^{2k} R(i_j,i_{j+1})\right)\right] = \sum_{i_1,i_2,\ldots,i_{2k}\in[1,n]} \mathbb{E}\left[\prod_{j=1}^{2k} R(i_j,i_{j+1})\right]$$

by linearity of expectation. Now, note that because $\mathbb{E}[R(i,j)] = 0$, the vast majority of the terms $\mathbb{E}[\prod_{j=1}^{2k} R(i_j,i_{j+1})]$ are 0; in fact, the only time the expected value is non-zero is when each consecutive pair of i's is distinct and when each $R(i,j)$ term appears an even number of times, in which case the expected value will be 1. Therefore, we can calculate the number of choices for $i_1, i_2, \ldots, i_{2k}$ that yield a non-zero value for $\mathbb{E}[\prod_{j=1}^{2k} R(i_j,i_{j+1})]$ and use that number to bound $\mathbb{E}[\mathrm{tr}((RR^T)^k)]$. We can think of the sum $\mathbb{E}[\prod_{j=1}^{2k} R(i_j,i_{j+1})]$ graphically as a sum over length $2k$ cycles in the vertex set $[1,n]$. We use what is known as a **constraint graph** to represent this cycle. In this case, the constraint graph consists of $2k$ vertices, each labeled from i_1 to i_{2k}; vertex i_j is connected to vertex i_{j+1} for all $1 \le j \le 2k$ to represent the term $R(i_j,i_{j+1})$, and a bold edge is drawn between i_r and i_s to signify that $i_r = i_s$.

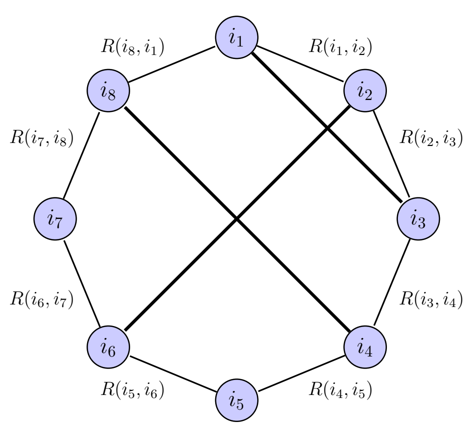

FIGURE 1. An example of a constraint graph where $k = 4$, $i_1 = i_3$, $i_2 = i_6$, and $i_4 = i_8$.

Note that in a scenario where j of $2k$ variables are equal, we only draw $j-1$ constraint edges to represent that equality, rather than $\binom{j}{2}$. This is because each constraint edge essentially represents a restriction; the extra constraint edges do not add to these restrictions, so they are not included.

Proposition 3.1. *In order for* $\mathbb{E}\big[\prod_{j=1}^{2k} R(i_j, i_{j+1})\big]$ *to have a non-zero value, there must be at least* $k-1$ *constraint edges in the respective constraint graph; in addition, this bound is sharp.*

Proof. We prove the first statement by induction on k. When $k=1$, the statement is vacuously true; $\mathbb{E}\big[\prod_{j=1}^{2k} R(i_j, i_{j+1})\big] = \mathbb{E}[R(i_1, i_2)^2]$, which has a non-zero value regardless of constraint edges.

Now, assume that the statement is true for $k = r$, and consider $k = r+1$. Assume $\mathbb{E}\big[\prod_{j=1}^{2k} R(i_j, i_{j+1})\big] \neq 0$, and consider the constraint graph. If each vertex is adjacent to at least one constraint edge, then because each constraint edge is adjacent to two vertices, there are at least $\frac{2r+2}{2} = r+1$ constraint edges, and we are done. Therefore, we only need to consider the case where there exists a vertex that is not adjacent to any constraint edges. Call this vertex i_j. Then, note that the statement $i_{j-1} = i_{j+1}$ must be true; if it was not, then the values $R(i_{j-1}, i_j)$ and $R(i_j, i_{j+1})$ have no corresponding equal terms, which means $\mathbb{E}\big[\prod_{j=1}^{2k} R(i_j, i_{j+1})\big] = 0$. But if $i_{j-1} = i_{j+1}$, then $R(i_{j-1}, i_j) = R(i_j, i_{j+1})$, meaning that we no longer need to consider the vertex i_j and its adjacent edges. Therefore, we can treat the vertices i_{j-1} and i_{j+1} as the same vertex, as they are equal, meaning that we have essentially reduced the constraint graph to one on $2r$ vertices. Then, by our Induction Hypothesis, this constraint graph requires at least $r-1$ constraint edges to create a nonzero expected value, which means that our total constraint graph requires at least r constraint edges, completing the proof.

In order to prove the sharpness of the bound, simply consider the case where $i_j = i_{2k+2-j}$ for all $2 \leq j \leq k$. Then, $R(i_l, i_{l+1}) = R(i_{2k+1-l}, i_{2k+2-l})$ for all $1 \leq l \leq k$, which creates a non-zero expected value. ∎

Now, we utilize Proposition 3.1 to bound the maximum number of times that $\mathbb{E}\big[\prod_{j=1}^{2k} R(i_j, i_{j+1})\big]$ can take a non-zero value, and use that information to bound $\mathbb{E}[\mathrm{tr}(R^{2k})]$.

Proposition 3.2. *Given a constraint graph on* b *vertices such that at least* c *constraint edges are required to create a non-zero expectation value, where each vertex has* n *possible values, let* N *represent the number of choices for the* b *vertices such that the expectation value of the product is non-zero. Then,* $N \leq \binom{b}{c} n^{b-c}(b-c)^c \leq b^{2c} n^{b-c}$.

Proof. Treat the set of vertices as an ordered set $S = \{d_1, d_2, \ldots, d_b\}$.

Because there must be at least c constraint edges, there must be at least c elements of S that are duplicates of other elements, so we can choose a set $I \subseteq S_b$ of c indices such that for all $j \in I$, there exists $m \notin I$ such that $d_j = d_m$. There are $\binom{b}{c}$ choices for I. We can then choose the elements $\{d_j \mid j \notin I\}$. Each element has at most n possible values so there are at most n^{b-c} choices for these elements. Finally, we choose the elements $\{d_j \mid j \in I\}$. To determine each d_j it is enough to specify the $m \notin I$ such that $d_j = d_m$. Each such d_j has $b-c$ choices, so there are at most $(b-c)^c$ choices for these elements. Therefore, $N \leq \binom{b}{c} n^{b-c}(b-c)^c$.

Now, note that $\binom{b}{c} \leq b^c$, as $\binom{b}{c} = \frac{b!}{(b-c)!c!} \leq \frac{b!}{(b-c)!} \leq b^c$. As $(b-c)^c \leq b^c$, this completes the proof. □

Corollary 3.3. *Let N represent the number of choices for the variables $(i_1, i_2, \ldots, i_{2k})$ such that $\mathbb{E}\big[\prod_{j=1}^{2k} R(i_j, i_{j+1})\big] \neq 0$. Then, $N \leq (2k)^{2k-2} n^{k+1}$.*

Proof. Apply Proposition 3.2. Note that $b = 2k$ and $c = k-1$ by Proposition 3.1. This implies the desired result.

□

Corollary 3.4. $\mathbb{E}[\mathrm{tr}(R^{2k})] \leq (2k)^{2k-2} n^{k+1}$.

Proof. Recall $\mathbb{E}[\mathrm{tr}(R^{2k})] = \sum_{i_1, i_2, \ldots, i_{2k} \in [1,n]} \mathbb{E}\Big[\prod_{j=1}^{2k} R(i_j, i_{j+1})\Big]$. By Corollary 3.3, the number of choices for $(i_1, i_2, \ldots, i_{2k})$ that yield a non-zero value for $\mathbb{E}\big[\prod_{j=1}^{2k} R(i_j, i_{j+1})\big]$ is at most $(2k)^{2k-2} n^{k+1}$; in addition, $\mathbb{E}\big[\prod_{j=1}^{2k} R(i_j, i_{j+1})\big] \leq 1$ for all choices of $(i_1, i_2, \ldots, i_{2k})$. These two observations complete the proof. □

Now, note that for any matrix R, $\mathrm{tr}(R^{2k})$ must take on a nonnegative value. Then, by Markov's inequality and Proposition 3.4,

$$\mathbb{P}[\mathrm{tr}(R^{2k}) \geq \frac{\mathbb{E}[\mathrm{tr}(R^{2k})]}{\epsilon}] \leq \epsilon \Longrightarrow \mathbb{P}[\mathrm{tr}(R^{2k}) \geq (2k)^{2k-2} n^{k+1}/\epsilon] \leq \epsilon.$$

This leads us to our main theorem for singular locally random matrices:

Theorem 3.5. *Given a random graph $G = G(n, \frac{1}{2})$ and R its singular locally random matrix, for all $\epsilon \in (0,1)$,*

$$\mathbb{P}[||R|| \geq e\sqrt{n}(\log(n/\epsilon) + 2)] \leq \epsilon.$$

Proof. Note that $||R|| \leq \sqrt[2k]{\mathrm{tr}((RR^T)^k)} = \sqrt[2k]{\mathrm{tr}(R^{2k})}$ for all positive integer k. In addition,

$$\mathbb{P}[\mathrm{tr}(R^{2k}) \geq (2k)^{2k-2} n^{k+1}/\epsilon] \leq \epsilon \Longrightarrow \mathbb{P}[\sqrt[2k]{\mathrm{tr}(R^{2k})} \geq \sqrt[2k]{(2k)^{2k-2} n^{k+1}/\epsilon}] \leq \epsilon.$$

But $\sqrt[2k]{(2k)^{2k-2}n^{k+1}/\epsilon} \leq \sqrt[2k]{(2k)^{2k}n^{k+1}/\epsilon} = 2k\sqrt{n}(n/\epsilon)^{1/2k}$. Setting $k = \lceil \log(n/\epsilon)/2 \rceil$, we see that $\sqrt[2k]{(2k)^{2k-2}n^{k+1}/\epsilon} \leq e\sqrt{n}(\log(n/\epsilon)+2)$.

Therefore, $\mathbb{P}[\sqrt[2k]{\mathrm{tr}(R^{2k})} \geq e\sqrt{n}(\log(n/\epsilon)+2)] \leq \epsilon$. As $||R|| \leq \sqrt[2k]{\mathrm{tr}(R^{2k})}$, the claim follows. $\square$

The important part of Theorem 3.5 is noticing that $||R||$ is $O(\sqrt{n}\log(n))$. Though it is known that $||R||$ is $O(\sqrt{n})$, the methods utilized in this section can be modified to yield similar results in the following sections.

4. Bounding the Norm of a Simple Locally Random Matrix

In this section, we generalize the technique utilized in the previous section to create a bound for all matrices of a certain form.

Definition 2. Consider a bipartite graph H with partite sets $X = \{x_1, x_2, \ldots, x_{|X|}\}$ and $Y = \{y_1, y_2, \ldots, y_{|Y|}\}$. Then, given a graph $(V(G), E(G)) = G = G(n, \frac{1}{2})$ with $|V| = n$, create a $\frac{n!}{(n-|X|)!} \times \frac{n!}{(n-|Y|)!}$ matrix R, known as a simple locally random matrix, with rows indexed by ordered sets of $|X|$ distinct numbers from 1 to n, and columns indexed by ordered sets of $|Y|$ distinct numbers from 1 to n. Then, if $A = \{a_1, a_2, \ldots, a_{|X|}\}$ and $B = \{b_1, b_2, \ldots, b_{|Y|}\}$, the entry of R indexed by A and B is defined as

$$R(A,B) = \begin{cases} 0 & A \cap B \neq \emptyset \\ (-1)^{E(A,B)} & A \cap B = \emptyset \end{cases}$$

where $E(A, B)$ is the number of pairs (i, j) such that edge (x_i, y_j) exists in H but edge (a_i, b_j) does not exist in G.

Note that the singular locally random matrix is obtained when H is just K_2. In addition, note that $\mathbb{E}[R(A, B)] = 0$ for randomly chosen A, B. Say $|X| = u$ and $|Y| = v$. Now, in order to find a probabilistic bound for $||R||$, we instead bound the norm of a closely related matrix.

Let $m = u + v$, and set $V_1, V_2, \ldots, V_m$ to be a partition of $\{1, 2, \ldots, n\}$ into m disjoint sets. Note that the total number of choices for the sets V is m^n, as each element of $\{1, 2, \ldots, n\}$ can belong to one of m possible partition sets. We can then define a matrix closely related to R:

Definition 3. Given $A = \{a_1, a_2, \ldots, a_u\}$ and $B = \{b_1, b_2, \ldots, b_v\}$ subsets of $\{1, 2, \ldots, n\}$, let $R_{V_1,\ldots,V_m}(A, B)$ be a $\frac{n!}{(n-u)!} \times \frac{n!}{(n-v)!}$ matrix in which $R_{V_1,\ldots,V_m}(A, B) = R(A, B)$ if $a_i \in V_i$ for all $i \leq u$ and $b_j \in V_{u+j}$ for all $j \leq v$; otherwise, $R_{V_1,\ldots,V_m}(A, B) = 0$.

Denote $R_{V_1,\ldots,V_m}$ as R'. We bound $||R'||$ by bounding $\mathbb{E}[\sqrt[2k]{\mathrm{tr}((R'R'^T)^k)}]$ and then use this information to bound $||R||$. Consider the bipartite graph H. Define $S_{n,u}$ to be the set of all sets of u distinct numbers chosen from 1 to n and define $S_{n,v}$ similarly. Then, notice

$$\begin{aligned}\mathbb{E}[\mathrm{tr}((R'R'^T)^k)] &= \mathbb{E}\left[\sum_{\substack{A_1,A_3,\ldots,A_{2k-1}\in S_{n,u}\\ B_2,B_4,\ldots,B_{2k}\in S_{n,v}}}\left(\prod_{j=1}^{k} R'(A_{2j-1},B_{2j})R'^T(B_{2j},A_{2j+1})\right)\right]\\ &= \sum_{\substack{A_1,A_3,\ldots,A_{2k-1}\in S_{n,u}\\ B_2,B_4,\ldots,B_{2k}\in S_{n,v}}}\mathbb{E}\left[\prod_{j=1}^{k} R'(A_{2j-1},B_{2j})R'^T(B_{2j},A_{2j+1})\right]\end{aligned}$$

by linearity of expectation. Denote $\prod_{j=1}^{k} R'(A_{2j-1},B_{2j})R'^T(B_{2j},A_{2j+1})$ as $P(A_1,\ldots,B_{2k})$. Similarly to the previous case, because $\mathbb{E}[R'(A,B)] = 0$, the vast majority of the terms $\mathbb{E}[P(A_1,\ldots,B_{2k})]$ are 0; the only time the expected value can be non-zero is when each consecutive pair of A's and B's is disjoint and every edge of G involved in the product appears an even number of times. In this case, the expected value will be at most 1. So, we can calculate the number of choices for $A_1, B_2, \ldots, A_{2k-1}, B_{2k}$ that yield a non-zero value for $\mathbb{E}[P(A_1,\ldots,B_{2k})]$ and use that number to bound $\mathbb{E}[\mathrm{tr}((R'R'^T)^k)]$. In order to represent $\mathbb{E}[P(A_1,\ldots,B_{2k})]$, we use another constraint graph.

This constraint graph is similar to the one in Proposition 3.1. In this constraint graph, there are $k(u+v)$ vertices sorted into $2k$ sets. These vertices are labeled $A_1 = \{a_{1;1}, a_{2;1}, \ldots, a_{u;1}\}$, $B_2 = \{b_{1;2}, b_{2;2}, \ldots, b_{v;2}\}$, $A_3 = \{a_{1;3}, a_{2;3}, \ldots, a_{u;3}\}, \ldots, B_{2k} = \{b_{1;2k}, b_{2;2k} \ldots, b_{v;2k}\}$. Two vertices $a_{p;q}$ and $b_{r;s}$ are adjacent in the constraint graph if and only if $|q-s| = 1$ and x_p and y_r are adjacent in H, where $a_{p;1} = a_{p;2k+1}$.

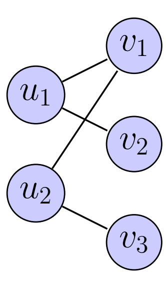

Figure 2a: H.

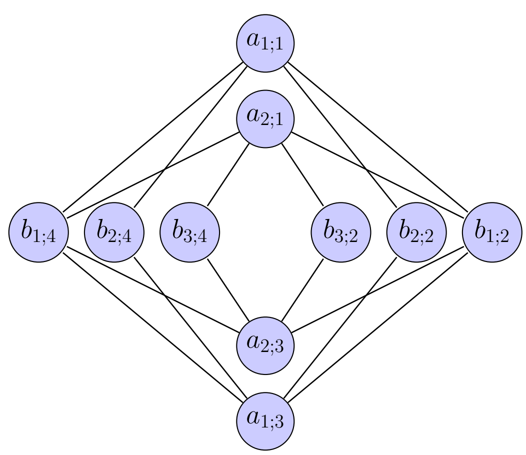

Figure 2b: An example of the constraint graph for the given example of H, where $k = 2$.

Now, in order to bound the number of choices for $A_1, B_2, \ldots, A_{2k-1}, B_{2k}$ that yield a non-zero expectation value, we can introduce the constraint edges again. However, note that due to the definition of R', constraint edges can only exist between vertices of the constraint graph that are created by the same vertex of H, as it is impossible for two vertices that are not created by the same vertex of H to be equal, as they correspond to different disjoint sets V_i, and the value of the variable must be in its respective set.

Lemma 4.1. *In order for $\mathbb{E}[P(A_1, \ldots, B_{2k})]$ to have a non-zero value, there must be at least $q(k-1)$ constraint edges in the respective constraint graph, where q is the size of a minimal vertex cover of H; in addition, this bound is sharp.*

Proof. In order to prove this proposition, we first show that the given bound is an upper bound then show that it is sharp by König's Theorem [19].

First, note that in order for $\mathbb{E}[P(A_1, \ldots, B_{2k})]$ to have a non-zero value, every edge in the constraint graph must have an equal counterpart by virtue of the constraint edges; this ensures that any edge that appears in the product appears an even number of times, creating a non-zero expected value.

It is easy to see that at most $q(k-1)$ constraint edges are required; namely, if V is a minimal vertex cover of H, then if $x_i \in V$, set $a_{i;1} = a_{i;3} = \cdots = a_{i;2k-1}$, and if $y_j \in V$, set $b_{j;2} = b_{j;4} = \cdots = b_{j;2k}$. Each such set of equalities corresponds to $k-1$ constraint edges, meaning that there are $q(k-1)$ constraint edges total. In addition, every edge in the constraint graph will have an equal counterpart by this method. If $(x_i, y_j) \in H$, at least one of x_i and y_j is in V by definition; without loss of generality $y_j \in V$. Then, this implies that for the edges in the constraint graph of the form $(a_{i;1}, b_{j;2}), (b_{j;2}, a_{i;3}), (a_{i;3}, b_{j;4}), \ldots, (b_{j;2k}, a_{i;1})$, each edge $(a_{i;2m-1}, b_{j,2m-2})$ has the equal counterpart $(a_{i;2m-1}, b_{j,2m})$ for $1 \leq m \leq k$, as $b_{j,2m-2} = b_{j,2m}$. Therefore, because each edge in the constraint graph is created by some edge in H, and all the edges in the constraint graph created by a certain edge in H are matched with an equal edge, this implies that this set of constraint edges is enough to create a non-zero expected value.

Now, we must show at least $q(k-1)$ constraint edges are required. Because H is a bipartite graph, we can apply König's Theorem, which states that there exists a matching of size q in H. Consider the q disjoint cycles of length $2k$ in the constraint graph that are created by the q edges in the matching of H. Because R' is defined so that constraint edges can only exist between vertices $a_{p;q}$ and $a_{p;q'}$ or between $b_{r;s}$ and $b_{r;s'}$, as two vertices not of this form do not belong to the same set V_i, any constraint edge created can affect at most 1 of the q cycles, due to the fact that all the cycles are disjoint and thus are impossible to link with a constraint edge. Therefore, each cycle requires at least $k-1$ constraint edges by Proposition 3.1, implying that the q cycles require at least $q(k-1)$ constraint edges total, completing the proof.

∎

Corollary 4.2. *Let N represent the number of choices for the sets $A_1, B_2, \ldots, A_{2k-1}, B_{2k}$ such that $\mathbb{E}[P(A_1, \ldots, B_{2k})] \neq 0$. Then, $N \leq ((u+v)k)^{2(k-1)q} n^{(u+v-q)k+q}$.*

Proof. Apply Proposition 3.2. In this situation, $b = k(u+v)$ and $c = q(k-1)$. This implies the desired result.

∎

Corollary 4.3. $\mathbb{E}[\mathrm{tr}((R'R'^T)^k)] \leq ((u+v)k)^{2(k-1)q} n^{(u+v-q)k+q}$.

Proof. Recall that $\mathbb{E}[\mathrm{tr}((R'R'^T)^k)] = \sum_{\substack{A_1, A_3, \ldots A_{2k-1} \in S_{n,l} \\ B_2, B_4, \ldots B_{2k} \in S_{n,m}}} \mathbb{E}\left[\prod_{j=1}^{k} R'(A_{2j-1}, B_{2j}) R'^T(B_{2j}, A_{2j+1})\right]$. Then, by Proposition 4.2, the number of choices for $A_1, B_2, \ldots A_{2k-1}, B_{2k}$ that yield a non-zero value for $\mathbb{E}\left[\prod_{j=1}^{k} R'(A_{2j-1}, B_{2j}) R'^T(B_{2j}, A_{2j+1})\right]$ is at most $((u+v)k)^{2(k-1)q} n^{(u+v-q)k+q}$; in addition, $\mathbb{E}\left[\prod_{j=1}^{k} R'(A_{2j-1}, B_{2j}) R'^T(B_{2j}, A_{2j+1})\right] \leq 1$. These two observations complete the proof. ∎

Now, note that for any graph G on n vertices, $\mathrm{tr}((R'R'^T)^k)$ must take on a nonnegative value. Then, by Markov's inequality and Corollary 4.3,

$$\mathbb{P}[\mathrm{tr}((R'R'^T)^k) \geq \frac{\mathbb{E}[\mathrm{tr}((R'R'^T)^k)]}{\epsilon}] \leq \epsilon \Longrightarrow \mathbb{P}[\mathrm{tr}((R'R'^T)^k) \geq ((u+v)k)^{2(k-1)q} n^{(u+v-q)k+q}/\epsilon] \leq \epsilon.$$

This allows us to bound the norm of our special matrix.

Proposition 4.4. *Given a random graph $G = G(n, \frac{1}{2})$, R its simple locally random matrix created by a bipartite graph H with partite sets of size u and v and minimal vertex cover of size q, and R' the matrix created by partitioning the row and column elements of R, then if $q \geq 1$ and $n \geq e^2$,*

$$\mathbb{P}[||R'|| \geq e^q n^{\frac{u+v-q}{2}} ((u+v)\log(n^q/\epsilon))^q] \leq \epsilon.$$

Proof. Note that $||R'|| \leq \sqrt[2k]{\mathrm{tr}((R'R'^T)^k)}$ for all positive integer k. In addition,

$$\mathbb{P}[\mathrm{tr}((R'R'^T)^k) \geq ((u+v)k)^{2(k-1)q} n^{(u+v-q)k+q}/\epsilon] \leq \epsilon \Longrightarrow$$

$$\mathbb{P}[\sqrt[2k]{\operatorname{tr}((R'R'^T)^k)} \geq \sqrt[2k]{((u+v)k)^{2(k-1)q}n^{(u+v-q)k+q}/\epsilon}] \leq \epsilon.$$

But $\sqrt[2k]{((u+v)k)^{2(k-1)q}n^{(u+v-q)k+q}/\epsilon} \leq ((u+v)k)^q n^{\frac{u+v-q}{2}}(n^q/\epsilon)^{1/2k}$.

Setting $k = \lceil \frac{1}{2q}\log(n^q/\epsilon)\rceil$ we find that $((u+v)k)^q n^{\frac{u+v-q}{2}}(n^q/\epsilon)^{1/2k} \leq e^q n^{\frac{u+v-q}{2}}((u+v)\log(n^q/\epsilon))^q$.

Therefore, $\mathbb{P}[\sqrt[2k]{\operatorname{tr}((R'R'^T)^k)} \geq e^q n^{\frac{u+v-q}{2}}((u+v)\log(n^q/\epsilon))^q] \leq \epsilon$. As $||R'|| \leq \sqrt[2k]{\operatorname{tr}((R'R'^T)^k)}$, the claim follows. $\square$

We can now use our bounds for $||R'||$ to bound $||R||$ through the following lemma.

Lemma 4.5. *Let M be a matrix and B, p be positive numbers such that*

(1) $M = \frac{1}{N}\sum_{V_1,\cdots,V_m} M_{V_1,\cdots,V_m}$ *for some matrices* $\{M_{V_1,\cdots,V_m}\}$ *where N is the number of possible* $V_1, \cdots, V_m$

(2) *For each choice of* $V_1, \cdots, V_m$*, for all* $x \in [\frac{1}{2}, N]$, $\mathbb{P}(||M_{V_1,\cdots,V_m}|| > Bx) \leq \frac{p}{64x^3}$

then $\mathbb{P}(||M|| \geq B) < p$.

Proof. The result follows from the following proposition.

Proposition 4.6. *For all $j \in [0, \lg N]$, the probability that there are more than $\frac{N}{2^{2j+2}}$ matrices $M_{V_1,\cdots,V_m}$ such that $||M_{V_1,\cdots,V_m}|| > 2^{j-1}B$ is at most $\frac{p}{2^{j+1}}$.*

Proof. We prove this by contradiction. If the probability that there are more than $\frac{N}{2^{2j+2}}$ matrices $M_{V_1,\cdots,V_m}$ such that $||M_{V_1,\cdots,V_m}|| > 2^{j-1}B$ is greater than $\frac{p}{2^{j+1}}$ then the probability that $||M_{V_1,\cdots,V_m}|| > 2^{j-1}B$ must be greater than $\frac{p}{2^{3j+3}}$. Plugging in $x = 2^{j-1}$, this gives a contradiction. $\square$

Using this proposition, with probability at least $1 - \sum_{j=0}^{\lfloor \lg n\rfloor}\frac{p}{2^{j+1}}$, for all integers j such that $0 \leq j \leq \lg N$, there are at most $\frac{N}{2^{2j+2}}$ matrices $M_{V_1,\cdots,V_m}$ such that $||M_{V_1,\cdots,V_m}|| > 2^{j-1}B$. When this occurs, for all integers j such that $0 \leq j \leq \lfloor \lg N\rfloor - 1$, there are at most $\frac{N}{2^{2j+2}}$ matrices $M_{V_1,\cdots,V_m}$ such that $2^{j-1}B < ||M_{V_1,\cdots,V_m}|| \leq 2^j B$. Moreover, there are no matrices such that $||M_{V_1,\cdots,V_m}|| > 2^{\lfloor \lg N\rfloor - 1}B$. This implies that with probability at least $1 - \sum_{j=0}^{\lfloor \lg n\rfloor}\frac{p}{2^{j+1}}$, $||M|| \leq \frac{B}{2} + \sum_{j=0}^{\lfloor \lg N\rfloor}\frac{2^j B}{2^{2j+2}} < B$, as needed. Since $1 - \sum_{j=0}^{\lfloor \lg n\rfloor}\frac{p}{2^{j+1}} > 1 - p$, the result follows. $\square$

Proposition 4.7. *Set $M = R$ and $M_{V_1,\ldots,V_m} = m^m R_{V_1,\ldots,V_m}$, where $q \geq 1$ and $n \geq e^3$. For all $p \in (0,1)$, let $B = (u+v)^{u+v}e^q n^{\frac{u+v-q}{2}}((u+v)\log(8n^q/p))^q$. Then, conditions* (1) *and* (2) *of Lemma 4.5 holds true for these values of M, B, and p.*

Proof. We must show both (1) and (2).

Note that $M = \frac{1}{N}\sum_{V_1,\dots,V_m} M_{V_1,\dots,V_m}$ because given a non-zero term $M(A,B)$, $R'(A,B)$ has probability $\frac{1}{(u+v)^{u+v}}$ of equaling $M(A,B)$ among all possible $V_1, V_2, \dots, V_m$; therefore, as $M_{V_1,\dots,V_m} = m^m R_{V_1,\dots,V_m}$, part (1) of the Lemma holds.

Now, we must show (2); that $\mathbb{P}[||M_{V_1,\dots,V_m}|| > Bx] = \mathbb{P}[||R'|| > Cx] \le \frac{p}{64x^3}$ for all $x \in [\frac{1}{2}, N]$, where $C = B/(u+v)^{u+v}$. If we let $f(x) = e^q n^{\frac{u+v-q}{2}}((u+v)\log(64n^q x^3/p))^q$, $P[||R'|| \ge f(x)] \le \frac{p}{64x^3}$ by Proposition 4.4. Define $g(x) = \frac{f(x)}{x}$. Note that $\lim_{x\to 0+} g(x) = -\infty$, and $\lim_{x\to\infty} g(x) = 0$. However, in the domain $x > 0$, $g'(x)$ has only one zero; at $x = \sqrt[3]{\frac{e^{3q}p}{64n^q}}$. As $n \ge e^3$, $x = \sqrt[3]{\frac{e^{3q}p}{64n^q}} \le \sqrt[3]{\frac{p}{64}} \le \sqrt[3]{\frac{1}{64}} = \frac{1}{4} < \frac{1}{2}$.

Therefore, for $x \ge \frac{1}{2}$, $g(x)$ is decreasing. But we know that $\mathbb{P}[||R'|| \ge f(x)] = \mathbb{P}[||R'|| \ge xg(x)] \le \frac{p}{64x^3}$. Therefore, as $C = g(1/2)$, for all $x \in [\frac{1}{2}, N]$, $\mathbb{P}[||R'|| \ge Cx] \le \mathbb{P}[||R'|| \ge xg(x)] \le \frac{p}{64x^3}$, as desired. □

Note that $((u+v)\log(8n^q/p))^q = (u+v)^q(\log(8n^q/p))^q$, and $e^q \le 3^{u+v+q}$. Therefore, $B = (u+v)^{u+v}e^q n^{\frac{u+v-q}{2}}((u+v)\log(8n^q/p))^q \le (3u+3v)^{u+v+q} n^{\frac{u+v-q}{2}} \log^q(8n^q/p)$. We use this bound to simplify the constants involved in the bound in our main theorem.

Now that we know our particular values of M, B, and p satisfy the conditions of Lemma 4.5, we apply the aforementioned lemma, leading us to our main theorem about simple locally random matrices:

Theorem 4.8. *Given a random graph $G = G(n, \frac{1}{2})$ and R its simple locally random matrix created by a bipartite graph H with partite sets of size u and v and minimal vertex cover of size q, if $q \ge 1$ and $n \ge e^3$, for all $\epsilon \in (0,1)$,*

$$\mathbb{P}[||R|| > (3u+3v)^{u+v+q} n^{\frac{u+v-q}{2}}(\log^q(8n^q/p))] < \epsilon.$$

In addition to an upper bound, our previous work allows us to calculate a simple lower bound for $\mathbb{E}[||R||]$.

Theorem 4.9. *Given a random graph $G = G(n, \frac{1}{2})$ and R its simple locally random matrix created by a bipartite graph H with partite sets of size u and v and minimal vertex cover of size q,*

$$\mathbb{E}[||R||] \ge \left(\frac{n}{m} - 1\right)^{\frac{u+v-q}{2}}.$$

Proof. The statement follows from the fact that $||R|| \ge ||R'||$ for all R', as R' is a matrix created by filling certain rows and columns of R with 0's. In addition, because we know Proposition 4.1

is sharp, there exists a constraint graph with $q(k-1)$ constraint edges that yields a non-zero expectation. Then we can pick $V_1, V_2, \ldots, V_m$ with $V_i = \{\lfloor \frac{n(i-1)}{m} \rfloor + 1, \lfloor \frac{n(i-1)}{m} \rfloor + 2, \ldots \lfloor \frac{n(i)}{m} \rfloor\}$ so that $\mathrm{tr}((R'R'^T)^k)$ is at least $(\frac{n}{m} - 1)^{(u+v)k-q(k-1)}$. As $\lim_{k\to\infty} \sqrt[2k]{\mathrm{tr}((R'R'^T)^k)} = ||R'||$, and $\lim_{k\to\infty} \sqrt[2k]{\mathrm{tr}((R'R'^T)^k)} \geq \lim_{k\to\infty} \sqrt[2k]{\left(\frac{n}{m} - 1\right)^{(u+v)k-q(k-1)}} = \left(\frac{n}{m} - 1\right)^{\frac{u+v-q}{2}}$, the claim follows.

□

Similarly to the singular locally random matrix case, the constant factors are not extremely relevant; it is more important to consider the asymptotic behavior of $||R||$ for large values of n. Therefore, the important part of Theorem 4.8 is noticing that $||R||$ is $O(n^{\frac{u+v-q}{2}} \log^q(n))$. In addition, Theorem 4.9 tells us that $\mathbb{E}[||R||]$ is $\Omega(n^{\frac{u+v-q}{2}})$, giving us relatively close upper and lower bounds for $||R||$.

5. Bounding the Norm of a Complex Locally Random Matrix

In this section, we consider norms of matrices that are defined using a much more complex graph for H. We utilize the same technique used in the past two sections to bound the norm of all matrices of this form. We first begin with a definition that allows us to rigorously define the type of matrices we are dealing with.

Definition 4. Given a graph $(V(H), E(H)) = H$ and a random labeled graph $(V(G), E(G)) = G(n, 1/2)$, let S be the set of all injective functions ϕ from $V(H)$ to $V(G)$. In addition, given a function $\phi \in S$, we then extend ϕ to be a function from $E(H)$ to possible edges of G by setting $\phi(h_1, h_2) = (\phi(h_1), \phi(h_2))$, where $h_1, h_2 \in V(H)$. Finally, we then define $\phi(E(H))$ to be the set of all possible edges in $E(G)$ that are mapped to from H by ϕ, and $\phi(E(H)) \setminus E(G)$ to be the set of possible edges in G that are mapped to by ϕ but do not exist in G.

Now that we have defined such a function ϕ, we can define our complex locally random matrices.

Definition 5. Consider a graph $(V(H), E(H)) = H$ with two disjoint subsets of $V(H)$ $X = \{x_1, x_2, \ldots x_{|X|}\}$ and $Y = \{y_1, y_2, \ldots y_{|Y|}\}$ such that no edges of H exist within X or within Y. Then, given a random graph $(V(G), E(G)) = G(n, 1/2)$, let M be the $\frac{n!}{(n-|X|)!} \times \frac{n!}{(n-|Y|)!}$ complex locally random matrix created by G and H, where the rows of M are indexed by ordered subsets of $\{1, 2, \ldots, n\}$ of size $|X|$ and the columns of M are indexed by ordered subsets of $\{1, 2, \ldots, n\}$ of size $|Y|$, with the entry of M indexed by the row $A = \{a_1, a_2, \ldots, a_{|X|}\}$ and the column $B = \{b_1, b_2, \ldots, b_{|Y|}\}$ being

$$\sum_{\phi}(-1)^{|\phi(E(H))\setminus E(G)|}$$

where the sum ranges over all injective functions ϕ from $V(H)$ to $V(G)$ such that $\phi(x_i) = a_i$ for all $1 \leq i \leq |X|$ and $\phi(y_j) = b_j$ for all $1 \leq j \leq |Y|$.

This type of matrix is much more complicated than simple locally random matrices; for simple locally random matrices, $X \cup Y = V(H)$. However, we can still apply the previous techniques with some modification in order to bound the norms of these matrices.

Denote $Z = V(H) \setminus X \setminus Y = \{z_1, z_2, \ldots, z_{|Z|}\}$, and say $|X| = u$, $|Y| = v$, and $|Z| = w$ for convenience. Now, in order to find a bound for $||M||$, we instead consider a closely related matrix.

Let $m = |V(H)| = u + v + w$, and set $V_1, V_2, \ldots, V_m$ to be a partition of $\{1, 2, \ldots, n\}$ into m disjoint sets. In addition, let θ to be a bijective function from $V(H)$ to $[1, m]$. Now, define $M_{V_1,\ldots,V_m}$ in the following manner:

Definition 6. Given our matrix M and $V_1, V_2, \ldots, V_m$ a partition of $\{1, 2, \ldots, n\}$ into m disjoint sets, then $M_{V_1,\ldots,V_m}$ is a $\frac{n!}{(n-u)!} \times \frac{n!}{(n-v)!}$ matrix with the rows indexed by ordered subsets of $\{1, 2, \ldots, n\}$ of size u and the columns indexed by ordered subsets of $\{1, 2, \ldots, n\}$ of size v, with the entry of $M_{V_1,\ldots,V_m}$ indexed by the row $\{a_1, a_2, \ldots, a_u\}$ and the column $\{b_1, b_2, \ldots, b_v\}$ being

$$\sum_{\phi'}(-1)^{|\phi'(E(H))\setminus E(G)|}$$

where this time the sum only ranges over injective functions ϕ' from $V(H)$ to $V(G)$ such that $\phi'(x_i) = a_i$ for all $1 \leq i \leq u$, $\phi'(y_j) = b_j$ for all $1 \leq j \leq v$, and $\phi'(h_1) \in V_{\theta(h_1)}$ for all $h_1 \in H$.

Note that in particular, this means that if there exists some i for which $a_i \notin V_{\theta(x_i)}$ or there exists some j for which $b_j \notin V_{\theta(y_j)}$, then the entry is 0.

Denote $M_{V_1,\ldots,V_m}$ as M' for convenience's sake. We can then define the entries of M' in a different method using the following definition and proposition.

Definition 7. Given three sets $A = \{a_1, a_2, \ldots, a_u\}$, $B = \{b_1, b_2, \ldots, b_v\}$, and $C = \{c_1, c_2, \ldots, c_w\}$, with each element of the sets in $[1, n]$, then define $Q(A, C, B) = (-1)^{|\phi'(E(H))\setminus E(G)|}$ where the function ϕ' is the unique injective function, if it exists, from $V(H)$ to $V(G)$ satisfying $\phi'(x_i) = a_i$ for all $1 \leq i \leq u$, $\phi'(y_j) = b_j$ for all $1 \leq j \leq v$, $\phi'(z_k) = c_k$ for all $1 \leq k \leq w$, and $\phi'(h_1) \in V_{\theta(h_1)}$ for all $h_1 \in H$. Note that if no such function ϕ exists, then $Q(A, C, B) = 0$. Finally, define $Q^T(B, C, A) = Q(A, C, B)$.

Utilizing this definition, we can find another way of calculating $M'(A,B)$. Substituting the ϕ' in Definition 6 with $Q(A,C,B)$, we can derive the following proposition.

Proposition 5.1. *Given an altered complex locally random matrix M', then the entry of M' indexed by the row $A = \{a_1, a_2, \ldots, a_u\}$ and the column $B = \{b_1, b_2, \ldots, b_v\}$ is*

$$\sum_{\substack{1\le c_1\le n\\ 1\le c_2\le n\\ \vdots\\ 1\le c_w\le n}} Q(A,C,B).$$

We now bound $||M'||$ and then use this information to bound $||M||$.

In order to find a probabilistic bound for $||M'||$, we bound we bound $\mathbb{E}[\sqrt[2k]{\mathrm{tr}((M'M'^T)^k)}]$. Define $S_{n,u}$ to be the set of all ordered sets of u distinct numbers chosen from 1 to n, and define $S_{n,v}$ similarly. Finally, let $S'_{n,w}$ be the set of all ordered sets of w numbers chosen from 1 to n, allowing repetition. Then, notice that

$$\begin{aligned}\mathbb{E}[\mathrm{tr}((M'M'^T)^k)] &= \sum_{\substack{A_1,A_5,\ldots,A_{4k-3}\in S_{n,u}\\ B_3,B_7,\ldots,B_{4k-1}\in S_{n,v}}} \mathbb{E}\left[\prod_{j=1}^{k} M'(A_{4j-3},B_{4j-1})M'^T(B_{4j-1},A_{4j+1})\right]\\ &= \sum_{\substack{A_1,A_5,\ldots,A_{4k-3}\in S_{n,u}\\ B_3,B_7,\ldots,B_{4k-1}\in S_{n,v}\\ C_2,C_4,\ldots C_{4k}\in S'_{n,w}}} \mathbb{E}\left[\prod_{j=1}^{k} Q(A_{4j-3},C_{4j-2},B_{4j-1})Q^T(B_{4j-1},C_{4j},A_{4j+1})\right]\end{aligned}$$

by linearity of expectation, where $A_{4k+1} = A_1$.

Denote $\prod_{j=1}^{k} Q(A_{4j-3},C_{4j-2},B_{4j-1})Q^T(B_{4j-1},C_{4j},A_{4j+1})$ as $P(A_1,C_2,\ldots,B_{4k-1},C_{4k})$. Similarly to the singular matrix case, because $\mathbb{E}[Q(A,C,B)] = 0$ for randomly chosen A, B, and C, the vast majority of the terms $\mathbb{E}[P(A_1,C_2,\ldots,B_{4k-1},C_{4k})]$ are 0; in fact, the only time the expected value can be non-zero is when each consecutive pair of sets of variables is disjoint and every edge of G involved in the product appears an even number of times, in which case the expected value is at most 1. So, we can calculate the number of choices for $A_1, C_2, \ldots, B_{4k-1}, C_{4k}$ that yield a non-zero value for $\mathbb{E}[P(A_1,C_2,\ldots,B_{4k-1},C_{4k})]$ and use that number to bound $\mathbb{E}[\mathrm{tr}((M'M'^T)^k)]$. In order to represent $\mathbb{E}[P(A_1,C_2,\ldots,B_{4k-1},C_{4k})]$, we use another constraint graph. This constraint graph is similar to the earlier one; however, it is slightly more complicated, as there is an extra set of vertices involved.

In this constraint graph, there are $k(u + 2w + v)$ vertices sorted into $4k$ sets. These vertices are labeled $A_1 = \{a_{1;1}, a_{2;1}, \ldots, a_{u;1}\}, C_2 = \{c_{1;2}, c_{2;2}, \ldots, c_{w;2}\}, B_3 = \{b_{1;3}, b_{2;3}, \ldots, b_{v;3}\}, C_4 = \{c_{1;4}, c_{2;4}, \ldots, c_{w;4}\}, A_5 = \{a_{1;5}, a_{2;5}, \ldots, a_{u;5}\}, \ldots, C_{4k} = \{c_{1;4k}, c_{2;4k} \ldots, c_{w;4k}\}$. Note that, in particular, the sets describing the sets of vertices are labeled A_1, C_2, B_3, C_4,…, and repeat this pattern exactly k times. Two vertices $a_{p;q}$ and $b_{r;s}$ are adjacent if and only if $|q - s| = 2$ and x_p and y_r are adjacent in H, where $a_{p;1} = a_{p;4k+1}$. Similarly, $a_{p;q}$ and $c_{t;o}$ are adjacent if and only if $|q - o| = 1$ and x_p and z_t are adjacent in H, and $b_{r;s}$ and $c_{t;o}$ are adjacent if and only if $|s - o| = 1$ and b_r and c_t are adjacent in H. Finally, $c_{t;o}$ and $c_{t';o'}$ are adjacent if and only if $o = o'$ and z_t and $z_{t'}$ are adjacent in H.

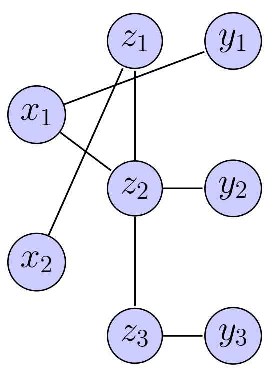

Figure 3a: H.

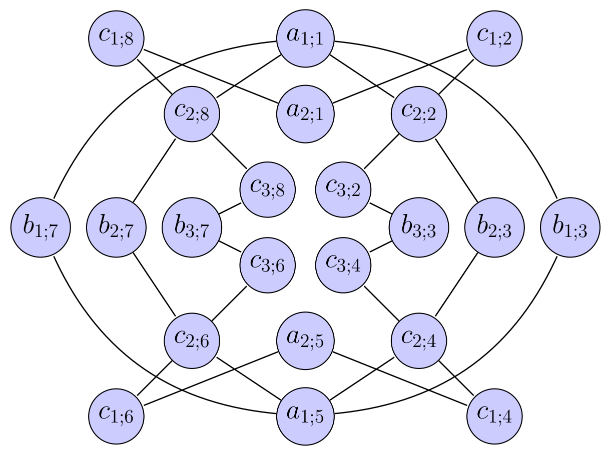

Figure 3b: An example of the constraint graph for the given example of H, where $k = 2$.

Now, in order to bound $\mathbb{E}[\text{tr}((M'M'^T)^k)]$, we must calculate the minimum number of constraint edges required in our constraint graph to bring about a non-zero expectation value, as that will help bound the number of choices for $A_1, C_2, B_3, C_4, A_5, \ldots, B_{4k-1}, C_{4k}$ that yield a non-zero expectation value for the product $\mathbb{E}[P(A_1, C_2, \ldots, B_{4k-1}, C_{4k})]$.

Lemma 5.2. *In order for $\mathbb{E}[P(A_1, C_2, \ldots, B_{4k-1}, C_{4k})]$ to have a non-zero value, there must be at least $q(k-1) + dk$ constraint edges in the respective constraint graph, where q is the maximal number of vertex-independent paths from X to Y in H and d is the number of vertices of C with positive degree; in addition, this bound is sharp.*

Proof. Consider q vertex-independent paths from X to Y. Create a multiset $S = \{s_1, s_2, \ldots s_q\}$ of size q including the sizes of all the vertex-independent paths from X to Y, where the size of a path is defined to be the number of edges in the path.

Note that any path of size s_i in H corresponds to a cycle of size $2ks_i$ in our constraint graph; this cycle requires $ks_i - 1$ dependencies by Proposition 3.1. Therefore, as the dependencies caused by two different cycles are entirely disjoint due to the cycles being disjoint, the total number of dependencies required by just the q vertex-independent paths is $\sum_{i=1}^{q}(ks_i - 1) = k(\sum_{i=1}^{q} s_i) - q$.

Consider the vertices in Z of positive degree but not in the q paths. Of the d vertices with positive degree in Z, because each pair of paths is disjoint and a path of length s_i corresponds to exactly $s_i - 1$ vertices in C, exactly $\sum_{i=1}^{q}(s_i - 1) = (\sum_{i=1}^{q} s_i) - q$ vertices of C are included in paths, so there are $d - ((\sum_{i=1}^{q} s_i) - q)$ vertices of Z with positive degree not included in the paths. Each of the vertices in Z of positive degree is adjacent to some edge in H, and that edge is repeated $2k$ times in the constraint graph; therefore, as each edge of G that appears in the constraint graph must appear an even number of times in the constraint graph, any particular vertex of positive degree can, in all of its appearances in the constraint graph, only take on at most k values. However, the particular vertex appears $2k$ times in the constraint graph, once in each set of vertices of the form C_i, so it must have at least k dependencies among those $2k$ appearances. Therefore, there are required to be a minimum of $(d - ((\sum_{i=1}^{q} s_i) - q))k + (k(\sum_{i=1}^{q} s_i) - q) = q(k-1) + dk$ dependencies in the constraint graph.

In order to prove the sharpness, we utilize Menger's Theorem [23]. First, note that the existence of q vertex-independent paths from X to Y implies that there exists q vertex-independent paths from X to Y, with first vertex in X, last vertex in Y, and all internal vertices in Z. This statement is true becausee given these q vertex-independent paths, if any of them have internal vertices in X or Y, we can simply shorten these paths until they have only first and last vertices in X and Y.

Now, Menger's Theorem states that because there are a maximum of q vertex-indepedent paths from X to Y in H with all internal vertices in Z, there is a set $S \in V(H)$ with $|S| = q$ such that all paths from X to Y pass through at least one vertex of S. Consider such a set S. Using S, we will create $q(k-1) + dk$ constraint edges that yield a non-zero expectation value for $\mathbb{E}[P(A_1, C_2, \ldots, B_{4k-1}, C_{4k})]$ as follows.

For all vertices $x_i \in X$ such that $x_i \in S$, set $a_{i;1} = a_{i;5} = \cdots = a_{i;4k-3}$. Note that this requires $k-1$ dependencies per vertex in S. Similarly, for all vertices $y_i \in Y$ such that $y_i \in S$, set $b_{i;3} = b_{i;7} = \cdots = b_{i;4k-1}$. In addition, for all vertices $z_i \in Z$ such that $z_i \in S$, set $c_{i;2} = c_{i;4} = \cdots = c_{i;4k}$. This requires $2k-1$ dependencies per vertex in S.

Now, consider all vertices $z_i \in Z$ such that $z_i \notin S$ and z_i has positive degree. Note that if there existed a path from z_i to X that passed through no vertices in S, there cannot exist a path from z_i to Y that passes through no vertices in S, as that would imply that there was a path from X to Y not passing through any vertices on S. So, if there exists a path from z_i to X passing through no vertices of S, set $c_{i;4} = c_{i;6}, c_{i;8} = c_{i;10}, \ldots, c_{i;4k} = c_{i;2}$. Otherwise, set $c_{i;2} = c_{i;4}, c_{i;6} = c_{i;8}, \ldots, c_{i;4k-2} = c_{i;4k}$. This requires k dependencies per vertex.

Now given an edge in the constraint graph $(a_{p;q}, b_{r;s})$ with $|q-s| = 2$, then either x_p or y_r is in S or else there would exist a path from X to Y not in S; therefore, either $(a_{p;q}, b_{r;q+2}) = (a_{p;q}, b_{r;q-2})$ or $(a_{p;s-2}, b_{r;s}) = (a_{p;s+2}, b_{r;s})$ by the constraint edges. Similarly, given $(a_{p;q}, c_{t;o})$ with $|p-o| = 1$, either $c_{t;q-1} = c_{t;q+1}$, which implies $(a_{p;q}, c_{t;q-1}) = (a_{p;q}, c_{t;q+1})$, or $z_t \notin S$ and $x_p \in S$, which implies $(a_{p;2o-q-2}, c_{t;2o-q-1}) = (a_{p;2o-q+2}, c_{t;2o-q+1})$. A similar argument applies to edges of the form $(c_{t;o}, b_{r;s})$. For edges of the form $(c_{t;o}, c_{t';o})$, it can be shown that either $(c_{t;o}, c_{t',o}) = (c_{t;o-2}, c_{t';o-2})$ or $(c_{t;o}, c_{t',o}) = (c_{t;o+2}, c_{t',o+2})$.

Finally, note that for if S contains exactly j vertices in Z, then the total number of constraint edges used in this method is $(k-1)(|S|-j) + (2k-1)j + (k)(d-j) = q(k-1) + dk$, meaning that the bound given is sharp.

□

Corollary 5.3. *Let N represent the number of choices for A_1, C_2, $\ldots$, B_{4k-1}, B_{4k} such that $\mathbb{E}[P(A_1, C_2, \ldots, B_{4k-1}, C_{4k})] \neq 0$. Then, $N \leq ((u+v+2w)k)^{2k(d+q)-2q} n^{(u+v+2w-d-q)k+q}$.*

Proof. Apply Proposition 3.2. In this situation, $b = k(u+v+2w)$ and $c = q(k-1) + dk$. This implies the desired result. □

Corollary 5.4. $\mathbb{E}[\mathrm{tr}((M'M'^T)^k)] \leq ((u+v+2w)k)^{2k(d+q)-2q} n^{(u+v+2w-d-q)k+q}$.

Proof.

Recall that $$\mathbb{E}[\mathrm{tr}((M'M'^T)^k)] = \sum_{\substack{A_1, A_5, \ldots, A_{4k-3} \in S_{n,u} \\ B_3, B_7, \ldots, B_{4k-1} \in S_{n,v} \\ C_2, C_4, \ldots C_{4k} \in S'_{n,w}}} \mathbb{E}\left[\prod_{j=1}^{k} Q(A_{4j-3}, C_{4j-2}, B_{4j-1}) Q^T(B_{4j-1}, C_{4j}, A_{4j+1})\right].$$
Then, by Proposition 5.3, there are at most $((u+v+2w)k)^{2k(d+q)-2q} n^{(u+v+2w-d-q)k+q}$ choices for $A_1, C_2, \ldots B_{4k-1}, C_{4k}$ that yield a non-zero value for $\mathbb{E}\left[\prod_{j=1}^{k} Q(A_{4j-3}, C_{4j-2}, B_{4j-1}) Q^T(B_{4j-1}, C_{4j}, A_{4j+1})\right]$; in addition, $\mathbb{E}\left[\prod_{j=1}^{k} Q(A_{4j-3}, C_{4j-2}, B_{4j-1}) Q^T(B_{4j-1}, C_{4j}, A_{4j+1})\right] \leq 1$. These two observations complete the proof. □

Now, note that for any graph G on n vertices, $\text{tr}((M'M'^T)^k)$ must take on a nonnegative value. Then, by Markov's Inequality and Corollary 5.4,

$$\mathbb{P}[\text{tr}((M'M'^T)^k) \geq ((u+v+2w)k)^{2k(d+q)-2q} n^{(u+v+2w-d-q)k+q}/\epsilon] \leq \epsilon.$$

This allows us to bound the norm of our special matrix.

Proposition 5.5. *Given a random graph $G = G(n, \frac{1}{2})$, M its complex locally random matrix created by H, and M' the matrix created by partitioning the row and column elements of M, then if $q \geq 1$ and $n \geq e^{2(d+q)}$,*

$$\mathbb{P}[||M'|| \geq e^{2(d+q)} n^{\frac{u+v+2w-d-q}{2}} ((u+v+2w)\log(n^q/\epsilon))^{d+q}] \leq \epsilon.$$

Proof. Note that $||M'|| \leq \sqrt[2k]{\text{tr}((M'M'^T)^k)}$ for all positive integer k. In addition,

$$\mathbb{P}[\text{tr}((M'M'^T)^k) \geq ((u+v+2w)k)^{2k(d+q)-2q} n^{(u+v+2w-d-q)k+q}/\epsilon] \leq \epsilon \implies$$

$$\mathbb{P}[\sqrt[2k]{\text{tr}((M'M'^T)^k)} \geq \sqrt[2k]{((u+v+2w)k)^{2k(d+q)-2q} n^{(u+v+2w-d-q)k+q}/\epsilon}] \leq \epsilon.$$

But $\sqrt[2k]{((u+v+2w)k)^{2k(d+q)-2q} n^{(u+v+2w-d-q)k+q}/\epsilon} \leq ((u+v+2w)k)^{d+q} n^{\frac{u+v+2w-d-q}{2}} (n^q/\epsilon)^{1/2k}$.

Setting $k = \lceil \frac{1}{2(d+q)} \log(n^q/\epsilon) \rceil$ we find that $((u+v+2w)k)^{d+q} n^{\frac{u+v+2w-d-q}{2}} (n^q/\epsilon)^{1/2k}$ is at most $e^{2(d+q)} n^{\frac{u+v+2w-d-q}{2}} ((u+v+2w)\log(n^q/\epsilon))^{d+q}$.

Therefore, $\mathbb{P}[\sqrt[2k]{\text{tr}((M'M'^T)^k)} \geq e^{2(d+q)} n^{\frac{u+v+2w-d-q}{2}} ((u+v+2w)\log(n^q/\epsilon))^{d+q}] \leq \epsilon$. As $||M'|| \leq \sqrt[2k]{\text{tr}((M'M'^T)^k)}$, the claim follows.

$\square$

Proposition 5.6. *Consider when $q \geq 1$ and $n \geq e^{3(d+q)}$. Set $Q = M$ and $Q_{V_1,\ldots,V_m} = m^m M_{V_1,\ldots,V_m}$. For all $p \in (0,1)$, let $B = (u+v+z)^{u+v+z} e^{2(d+q)} n^{\frac{u+v+2w-d-q}{2}} ((u+v+2w)\log(8n^q/p))^{d+q}$. Then Conditions* (1) *and* (2) *of Lemma 4.5 holds true for these values of Q, B, and p.*

Proof. We must show both (1) and (2).

Note that $Q = \frac{1}{N}\sum_{V_1,\ldots,V_m} Q_{V_1,\ldots,V_m}$ because given a non-zero term $Q(A,B)$, $M'(A,B)$ has probability $\frac{1}{(u+v+w)^{u+v+w}}$ of equaling $Q(A,B)$ among all possible $V_1, V_2, \ldots, V_m$; therefore, as $Q_{V_1,\ldots,V_m} = m^m M_{V_1,\ldots,V_m}$, part (1) of the Lemma holds.

Now, we must show (2); that $\mathbb{P}[||Q_{V_1,\ldots,V_m}|| > Bx] = \mathbb{P}[||M'|| > Cx] \leq \frac{p}{64x^3}$ for all $x \in [\frac{1}{2}, N]$, where $C = B/(u+v+w)^{u+v+w}$. If we let $f(x) = e^{2(d+q)} n^{\frac{u+v+2w-d-q}{2}} ((u+v+2w)\log(64x^3 n^q/p))^{d+q}$, $P[||M'|| \geq f(x)] \leq \frac{p}{64x^3}$ by Proposition 5.5. Define $g(x) = \frac{f(x)}{x}$. Note that $\lim\limits_{x\to 0+} g(x) = -\infty$, and

$\lim_{x\to\infty} g(x) = 0$. However, in the domain $x > 0$, $g'(x)$ has only one zero; at $x = \sqrt[3]{\frac{e^{3(d+q)}p}{64n^q}}$. As $n \geq e^{3(d+q)}$, $x = \sqrt[3]{\frac{e^{3(d+q)}p}{64n^q}} \leq \sqrt[3]{\frac{p}{64}} \leq \sqrt[3]{\frac{1}{64}} = \frac{1}{4} < \frac{1}{2}$.

Therefore, for $x \geq \frac{1}{2}$, $g(x)$ is decreasing. But we know that $\mathbb{P}[||M'|| \geq f(x)] = \mathbb{P}[||M'|| \geq xg(x)] \leq \frac{p}{64x^3}$. Therefore, as $C = g(1/2)$, for all $x \in [\frac{1}{2}, N]$, $\mathbb{P}[||M'|| \geq Cx] \leq \mathbb{P}[||M'|| \geq xg(x)] \leq \frac{p}{64x^3}$, as desired.

□

Note that $((u+v+2w)\log(8n^q/p))^{d+q} \leq (u+v+2w)^{u+v+w}(\log(8n^q/p))^{d+q}$, and $e^{d+q} \leq 3^{2u+2v+2w}$. Therefore, $B \leq (3u+3v+6w)^{2u+2v+2w} n^{\frac{u+v+2w-d-q}{2}} (\log(8n^q/p))^{d+q}$. We use this bound to simplify the constants involved in the bound in our main theorem.

Theorem 5.7. *for all $\epsilon \in (0,1)$,*

$$\mathbb{P}[||M|| > (3u+3v+6w)^{2u+2v+2w} n^{\frac{u+v+2w-d-q}{2}} (\log(8n^q/\epsilon))^{d+q}] < \epsilon.$$

In addition, our previous work allows us to calculate a simple lower bound for $\mathbb{E}[||M||]$.

Theorem 5.8. *Given a random graph $G = G(n, \frac{1}{2})$ and M its complex locally random matrix created by a graph H with partite sets of size u and v and middle set of size w, with a maximum of q vertex-independent paths from set X to set Y and with d middle vertices of non-zero degree, then if $m = u+v+w$,*

$$\mathbb{E}[||M||] \geq \left(\frac{n}{m} - 1\right)^{\frac{u+v+2w-d-q}{2}}.$$

Proof. The statement follows from the fact that $||M|| \geq ||M'||$ for all R', as R' is a matrix created by filling certain rows and columns of M with 0's. In addition, because we know Proposition 5.2 is sharp, there exists a constraint graph with $q(k-1)+dk$ constraint edges that yields a non-zero expectation. Then we can pick $V_1, V_2, \ldots, V_m$ with $V_i = \{\lfloor\frac{n(i-1)}{m}\rfloor + 1, \lfloor\frac{n(i-1)}{m}\rfloor + 2, \ldots \lfloor\frac{n(i)}{m}\rfloor\}$ so that $\mathrm{tr}((M'M'^T)^k)$ is at least $(\frac{n}{m} - 1)^{(u+v+2w)k-q(k-1)-dk}$. As $\lim_{k\to\infty} \sqrt[2k]{\mathrm{tr}((M'M'^T)^k)} = ||M'||$, and $\lim_{k\to\infty} \sqrt[2k]{\mathrm{tr}((M'M'^T)^k)} \geq \lim_{k\to\infty} \sqrt[2k]{\left(\frac{n}{m} - 1\right)^{(u+v+2w)k-q(k-1)-dk}} = \left(\frac{n}{m} - 1\right)^{\frac{u+v+2w-d-q}{2}}$, the claim follows.

□

Therefore, Theorem 5.7 tells us that $||M||$ is $O(n^{\frac{u+v+2w-d-q}{2}}(\log(n))^{d+q})$, and Theorem 5.8 tells us that $\mathbb{E}[||M||]$ is $\Omega(n^{\frac{u+v+2w-d-q}{2}})$. Together, these two statements allow us to find a relatively strong upper and lower bound for $||M||$.

6. Applications

Norm bounds on locally random matrices have been needed in proving lower bounds for the Sum of Squares Hierarchy for the planted clique problem several times. These bounds have previously been handled case by case. Theorems 4.8 and 5.7 are unified theorems that give norm bounds on all such matrices.

One example of the applications of this work is in [22]. Section 9 of this paper focuses on bounding the norm of an $\binom{n}{a} \times \binom{n}{a}$ locally random matrix R_a with entries defined as

$$R_a(V,W) = \begin{cases} 2^{a^2} - 1 & V \cap W = \emptyset, \{(v,w) : v \in V, w \in W\} \subseteq E(G) \\ -1 & V \cap W = \emptyset, \{(v,w) : v \in V, w \in W\} \nsubseteq E(G) \\ 0 & V \cap W \neq \emptyset \end{cases}$$

where V and W are subsets of $[1,n]$ of size a. Define R_H to be the simple locally random matrix created by a random graph $G = G(n, 1/2)$ and the bipartite graph H; then, create the matrix $R'_a = \sum_H R_H$, where the sum ranges over all bipartite graphs H with both partite sets of size a and $|E(H)| > 0$. Because $|E(H)| > 0$, it can be shown that $||R_H||$ is $O(n^{a-\frac{1}{2}} \log(n))$ for all H by Theorem 4.8. Because the norm is subadditive, this would imply that $||R'_a||$ is $O(n^{a-\frac{1}{2}} \log(n))$ as well. However, it can be shown that R'_a is just an extended version of R_a, in which every row and column is repeated $a!$ times; therefore, because R_a is a submatrix of R'_a, $||R_a||$ is $O(n^{a-\frac{1}{2}} \log(n))$, which is what a major part of [22] focuses on proving.

Similarly, Theorems 4.8 and 5.7 give a direct proof of probabilistic norm bounds shown in [17] and [27], though the details are omitted for sake of space.

7. Conclusion and Further Studies

In this paper, we analyzed the norms of certain matrices that were associated with random graphs $G = G(n, 1/2)$. We proved that the simple locally random matrix R created by such a graph G and a bipartite graph H with partite sets of size u and v and minimal vertex cover of size q had a norm that was $O(n^{\frac{u+v-q}{2}} \log^q(n))$ with high probability and had an expected value that was $\Omega(n^{\frac{u+v-w}{2}})$. More generally, we also showed that the complex locally random matrix M associated with G and a graph H with two sets A and B of size u and v had a norm of size $O(n^{\frac{u+v+2w-d-q}{2}}(\log(n))^{d+q})$ with high probability and had expected value that was $\Omega(n^{\frac{u+v+2w-d-q}{2}})$, where $w = |H| - u - v$, q is the maximum number of vertex-independent paths from A to B, and d is the number of vertices in $V(H) \setminus A \setminus B$ of positive degree.

As described in the previous section, special cases of these bounds have appeared in analyzing the performance of the Sum of Squares Hierarchy on the planted clique problem. The general bounds presented in this paper will very likely be useful for further analysis of the Sum of Squares's performance on the planted clique problem, which would allow one to more accurately calculate the power of the SOS Hierarchy and possibly determine a better algorithm for finding planted cliques. In addition, the bounds in this paper may be useful for analyzing the performance of the Sum of Squares Hierarchy on other problems as well.

A further research question is whether the norm bounds on these matrices can be tightened by removing the polylog factor from the upper bounds, which would result in a much stronger approximation for the norms of these matrices. In particular, better approximations in the trace method and Proposition 3.2 may help lower or possibly remove this polylog factor. In addition, it might be fruitful to consider locally random matrices with a different distribution for the entries.

8. Acknowledgements

The author thanks the Program for Research in Mathematics, Engineering, and Science (PRIMES-USA) at MIT for supporting this research. In addition, the author thanks Aaron Potechin, the mentor for this project, for suggesting this project and providing invaluable guidance along the way. Finally, the author thanks the PRIMES-USA faculty, including Dr. Tanya Khovanova, Dr. Pavel Etingof, and Dr. Slava Gerovitch, for making this research possible.

References

[1] Noga Alon, Alexandr Andoni, Tali Kaufman, Kevin Matulef, Ronitt Rubinfeld, and Ning Xie. Testing k-wise and almost k-wise independence. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 496–505. ACM, 2007.

[2] Noga Alon, Michael Krivelevich, and Benny Sudakov. Finding a large hidden clique in a random graph. *Random Structures and Algorithms*, 13(3-4):457–466, 1998.

[3] Benny Applebaum, Boaz Barak, and Avi Wigderson. Public-key cryptography from different assumptions. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 171–180. ACM, 2010.

[4] Sanjeev Arora, Boaz Barak, Markus Brunnermeier, and Rong Ge. Computational complexity and information asymmetry in financial products. In *ICS*, pages 49–65, 2010.

[5] Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for unique games and related problems. In *Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium on*, pages 563–572. IEEE, 2010.

[6] Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings and graph partitioning. *Journal of the ACM (JACM)*, 56(2):5, 2009.

[7] Z Bai and YQ Yin. Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. 1993.

[8] Boaz Barak, Fernando GSL Brandao, Aram W Harrow, Jonathan Kelner, David Steurer, and Yuan Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 307–326. ACM, 2012.

[9] Boaz Barak, Prasad Raghavendra, and David Steurer. Rounding semidefinite programming hierarchies via global correlation. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 472–481. IEEE, 2011.

[10] Quentin Berthet and Philippe Rigollet. Complexity theoretic lower bounds for sparse principal component detection. In *Conference on Learning Theory*, pages 1046–1066, 2013.

[11] Yael Dekel, Ori Gurel-Gurevich, and Yuval Peres. Finding hidden cliques in linear time with high probability. *Combinatorics, Probability and Computing*, 23(01):29–49, 2014.

[12] Yash Deshpande and Andrea Montanari. Improved sum-of-squares lower bounds for hidden clique and hidden submatrix problems. *CoRR*, abs/1502.06590, 2015.

[13] Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.

[14] Dima Grigoriev. Complexity of positivstellensatz proofs for the knapsack. *Computational Complexity*, 10(2):139–154, 2001.

[15] Dima Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1):613–622, 2001.

[16] Venkatesan Guruswami and Ali Kemal Sinop. Lasserre hierarchy, higher eigenvalues, and approximation schemes for graph partitioning and quadratic integer programming with psd objectives. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 482–491. IEEE, 2011.

[17] Samuel B. Hopkins, Pravesh K. Kothari, and Aaron Potechin. Sos and planted clique: Tight analysis of MPW moments at all degrees and an optimal lower bound at degree four. *CoRR*, abs/1507.05230, 2015.

[18] Mark Jerrum. Large cliques elude the metropolis process. *Random Structures & Algorithms*, 3(4):347–359, 1992.

[19] Denés Konig. Gráfok és mátrixok. matematikai és fizikai lapok, 38: 116–119, 1931.

[20] Luděk Kučera. Expected complexity of graph partitioning problems. *Discrete Applied Mathematics*, 57(2):193–212, 1995.

[21] Jean B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[22] R. Meka, A. Potechin, and A. Wigderson. Sum-of-squares lower bounds for planted clique. *ArXiv e-prints*, March 2015.

[23] Karl Menger. Zur allgemeinen kurventheorie. *Fundamenta Mathematicae*, 10(1):96–115, 1927.

[24] Yurii Nesterov. Squared functional systems and optimization problems. In *High performance optimization*, pages 405–440. Springer, 2000.

[25] Pablo A Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, Citeseer, 2000.

[26] Pavel A Pevzner, Sing-Hoi Sze, et al. Combinatorial approaches to finding subtle signals in dna sequences. In *ISMB*, volume 8, pages 269–278, 2000.

[27] Prasad Raghavendra and Tselil Schramm. Tight lower bounds for planted clique in the degree-4 SOS program. *CoRR*, abs/1507.05136, 2015.

[28] Grant Schoenebeck. Linear level lasserre lower bounds for certain k-csps. In *Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on*, pages 593–602. IEEE, 2008.

[29] NZ Shor. Class of global minimum bounds of polynomial functions. *Cybernetics and Systems Analysis*, 23(6):731–734, 1987.

[30] Eugene P Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, pages 325–327, 1958.

[31] Eugene P Wigner. Characteristic vectors of bordered matrices with infinite dimensions ii. In *The Collected Works of Eugene Paul Wigner*, pages 541–545. Springer, 1993.