Lecture Series 4

August 29, 2025

We now introduce the concept of graph sparsification, and see how we could apply this to Ramanujan graphs, mainly in reference to [3].

A dense graph is one where the number of edges is close to the maximum number of edges that could possibly exist. In other words, a dense graph has a relatively high edge-to-vertex ratio. A sparse graph is one where the number of edges is relatively small compared to the number of vertices. In other words, a sparse graph has a low edge-to-vertex ratio.

The goal of graph sparsification is to approximate a given graph $G = (V, E, \omega)$ by a sparse graph $H = (V, F, \hat{\omega})$ on the same set of vertices. In a sense, we want to create simpler version of the graph, retaining essential properties while having less edges to deal with, reducing the complexity of the graph. This can allow us to make approximations without having too much errors.

1 Graph Sparsification

There are various ways to define graph sparsification. In this chapter, we will say that H is a $\kappa-$approximation of G if for all $x \in \mathbb{R}^V$,

$$\boldsymbol{x}^T L_G \boldsymbol{x} \leq \boldsymbol{x}^T L_H \boldsymbol{x} \leq \kappa \cdot \boldsymbol{x}^T L_G \boldsymbol{x},$$

where L_G and L_H are the Laplacian matrices of G and H.

Note that previously in Chapter 1, we defined Laplacian matrix, Adjacency and Degree matrices in the unweighted setting. In a weighted graph $G = (V, E, w)$, the Laplacian L is still defined the same way by $L = D - A$. However, D and A are defined slightly differently, and is defined as follows:

$$A_{i,j} \stackrel{\text{def}}{=} \begin{cases} w_G(ij) & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}, D_{i,j} \stackrel{\text{def}}{=} \begin{cases} \sum_k w_G(ik) & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

We could also express $x^T Lx$ in a quadratic form, given by $x^T Lx = \sum_{\{u,v\} \in E} (x_u - x_v)^2 w_{uv}$.

In the case where G is the complete graph, good spectral sparsifiers are supplied by Ramanujan Graphs. These are $d-$regular graphs H all of whose non-zero Laplacian eigenvalues lie between $d - 2\sqrt{d-1}$ and $d + 2\sqrt{d-1}$. Hence if we take a Ramanujan graph on n vertices and multiply every edge by $n/(d-2\sqrt{d-1})$, we obtain a graph taht $\kappa-$approximates the complete graph, for

$$\kappa = \frac{d+2\sqrt{d-1}}{d-2\sqrt{d-1}}$$

We could show that every graph can be apprixmated at least this well by a graph with only twice as many edges as the Ramanujan graph, as a $d-$regular graph has $dn/2$ edges.

Theorem 1.1. *For every $d > 1$, every undirected weighted graph $G = (V, E, \omega)$ on n vertices contains a weighted subgraph $H = (V, F, \hat{\omega})$ with $\lceil d(n-1) \rceil$ edges that satisfies:*

$$\boldsymbol{x}^T L_G \boldsymbol{x} \leq \boldsymbol{x}^T L_H \boldsymbol{x} \leq \left(\frac{d+1+2\sqrt{d}}{d+1-2\sqrt{d}} \right) \cdot \boldsymbol{x}^T L_G \boldsymbol{x} \quad \forall \boldsymbol{x} \in \mathbb{R}^V.$$

1.1 The Incidence Matrix and the Laplacian

Given a connected weighted undirected graph $G = (V, E, w)$ with $w_e > 0$, if we orient the edges of G arbitrarily, we can write its Laplacian as $L = B^T W B$, where $B_{m\times n}$ is the signed edge-vertex incidence matrix, given by

$$B_{i,j} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if edge } e_j \text{ enters vertex } v_i \\ -1 & \text{if edge } e_j \text{ leaves vertex } v_i \\ 0 & \text{otherwise} \end{cases}$$

and $W_{m\times m}$ is the diagonal matrix with $W_{e,e} = w_e$.

We denote the row vectors of B by $\{b_e\}_{e\in E}$ and the span of its column by $\mathbb{B} = im(B) \subseteq \mathbb{R}^m$. Note that $b^T_{(u,v)} = (x_v - x_u)$.

Then L is positive semidefinite since for every $s \in \mathbb{R}^n$,

$$x^T L x = x^T B^T W B x = ||W^{1/2} B x||_2^2 \geq 0.$$

We also have $ker(L) = ker(W^{1/2}B) = span(\mathbf{1})$ since

$$\begin{aligned} x^T L x = 0 &\iff ||W^{1/2} B x||_2^2 \\ &\iff \sum_{(u,v)\in E} w_{(u,v)} (x_u - x_v)^2 = 0 \\ &\iff x_u - x_v = 0 \quad \forall (u,v) \\ &\iff x \text{ is constant, since } G \text{ is connected.} \end{aligned}$$

1.2 Moore-Penrose Pseudoinverse

The Moore-Penrose Pseudoinverse [4] is the generalization of the inverse matrix. The pseudoinverse generalizes the inverse as the matrix does not have to be invertible or a square matrix. We denote the pseudoinverse of the matrix A as A^+. The pseudoinverse is very important in the development of the applications and the Laplacian Systems. This is because while the Laplacian is a square matrix, it may not be invertible. Formally, we define the pseudoinverse as follows:

Definition 1.2. *For $A \in \mathbb{F}^{m\times n}$, where $\mathbb{F}$ is a field of either real or complex numbers, a pseudoinverse of A is defined as $A^+ \in \mathbb{F}^{n\times m}$ if it satisfy all of the four following criteria:*

1. $AA^+A = A$
2. $A^+AA^+ = A^+$
3. $(AA^+)^* = AA^+$
4. $(A^+A)^* = A^+A$

Since L is symmetric, we can diagonalize it and write

$$L = \sum_{i=2}^{n} \lambda_i u_i u_i^T$$

where $\lambda_2, ..., \lambda_n$ are the nonzero eigenvalues of L and $u_2, ..., u_n$ are the corresponding set of orthonormal eigenvectors. The Moore-Penrose Pseudoinverse of L is then defined as

$$L^+ = \sum_{i=2}^{n} \frac{1}{\lambda_i} u_i u_i^T.$$

Notice that $ker(L) = ker(L^+)$ and

$$LL^+ = L^+L = \sum_{i=2}^{n} u_i u_i^T,$$

which is the projection onto the span of nonzero eigenvectors of L and L^+. Hence $LL^+ = L^+L$ is the identity on $im(L) = ker(L)^\perp = span(\mathbf{1})^\perp$.

1.3 Rank-one updates

We now introduce a well known theorem in linear algebra, which describes the behaviour of the inverse of a matrix under rank-one updates. This is also known as the Sherman-Morisson Formula.

Lemma 1.3 (Sherman-Morisson Formula)**.** *[2] If A is a nonsungular $n \times n$ matrix and $\boldsymbol{v}$ is a vector, then*

$$(A + \boldsymbol{v}\boldsymbol{v}^T)^{-1} = A^{-1} - \frac{A^{-1}\boldsymbol{v}\boldsymbol{v}^T A^{-1}}{1 + \boldsymbol{v}^T A^{-1}\boldsymbol{v}}$$

A related formula describing the change in determinant is the following

Lemma 1.4 (Matrix Determinant Lemma)**.** *If A is nonsingular and $\boldsymbol{v}$ is a vector, then*

$$\det(A + \boldsymbol{v}\boldsymbol{v}^T) = \det(A)(1 + \boldsymbol{v}^T A^{-1}\boldsymbol{v}).$$

1.4 Proof of sparsification

In this subsection, we introduce the Theorem that proves Theorem 1.1. The proof of this Theorem requires several linear algebraic theorem. Here, we use $A \preceq B$ to mean that $B - A$ is positive semidefinite, and $\mathbf{id}_S$ denotes the identity operator on a vector space S.

Theorem 1.5. *Suppose d¿1 and $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_m$ are vectors in $\mathbb{R}^n$ with*

$$\sum_{i \leq m} \boldsymbol{v}_i \boldsymbol{v}_i^T = \mathbf{id}_{\mathbb{R}^n} .$$

Then there exist scalars $s_i \geq 0$ with $|\{i : s_i \neq 0\}| \leq dn$ such that

$$\mathbf{id}_{\mathbb{R}^n} \preceq \sum_{i \leq m} s_i \boldsymbol{v}_i \boldsymbol{v}_i^T \preceq \left(\frac{d + 1 + 2\sqrt{d}}{d + 1 - 2\sqrt{d}} \right) \mathbf{id}_{\mathbb{R}^n} .$$

Theorem 1.1 follows from this Theorem.

Proof of Theorem 1.1. Without loss of generality, assume G is connected. Write $L_G = B^T W B$ and fix $d > 1$. Restrict attention to $\mathrm{im}(L_G) \cong \mathbb{R}^{n-1}$ and apply Theorem 1.5 to the columns $\{\mathbf{v}_i\}_{i \leq m}$ of

$$V_{n \times m} = (L_G^+)^{\frac{1}{2}} B^T W^{\frac{1}{2}}$$

which are indexed by the edges of G and satisfy

$$\begin{aligned} \sum_{i \leq m} \mathbf{v}_i \mathbf{v}_i^T = VV^T &= (L_G^+)^{\frac{1}{2}} B^T W B^T (L_G^+)^{\frac{1}{2}} \\ &= (L_G^+)^{\frac{1}{2}} L_G^+ (L_G^+)^{\frac{1}{2}} \\ &= \mathbf{id}_{\mathrm{im}(L_G)} \end{aligned}$$

Write the scalars $s_i \geq 0$ guaranteed by the theorem in the $m \times m$ diagonal matrix $S(i, i) = s_i$ and set $L_H = B^T W^{\frac{1}{2}} S W^{\frac{1}{2}} B$. Then L_H is the Laplacian of the subgraph H of G with edge weights $\{\tilde{w}_i = w_i s_i\}_{i \in E}$, and H has at most $d(n-1)$ edges since at most that many of the s_i are nonzero. Also,

$$\mathbf{id}_{\mathrm{im}(L_G)} \preceq \sum_{i \leq m} s_i \mathbf{v}_i \mathbf{v}_i^T = VSV^T \preceq \kappa \cdot \mathbf{id}_{\mathrm{im}(L_G)} \quad \text{for } \kappa = \frac{d + 1 + 2\sqrt{d}}{d + 1 - 2\sqrt{d}}$$

Then by variational characterization of eigenvalues, we have

$$\begin{aligned}
& 1 \leq \frac{y^T V S V^T y}{y^T y} \leq \kappa \quad \forall y \in \operatorname{im}((L_G)^{\frac{1}{2}}) = \operatorname{im}(L_G) \\
\iff & 1 \leq \frac{y^T (L_G^+)^{\frac{1}{2}} L_H (L_G^+)^{\frac{1}{2}} y}{y^T y} \leq \kappa \quad \forall y \in \operatorname{im}((L_G)^{\frac{1}{2}}) \\
\iff & 1 \leq \frac{x^T L_G^{\frac{1}{2}} (L_G^+)^{\frac{1}{2}} L_H (L_G^+)^{\frac{1}{2}} L_G^{\frac{1}{2}} x}{x^T L_G^{\frac{1}{2}} L_G^{\frac{1}{2}} x} \leq \kappa \quad \forall x \perp \mathbf{1} \\
\iff & 1 \leq \frac{x^T L_H x}{x^T L_G x} \leq \kappa \quad \forall x \perp \mathbf{1}
\end{aligned}$$

□

Now, we will build up the proof of Theorem 1.5. Theorem 1.5 is well-linked to the Kadison-Singer conjecture which we will introduce later.

1.5 Proving Theorem 1.5

1.5.1 An Intuition

It is well known that the eigenvalues of $A + \mathbf{v}\mathbf{v}^T$ interlace those of A. In fact the new eigenvalues can be determined exactly by looking at the characteristic polynomial of $A + \mathbf{v}\mathbf{v}^T$ which is computed by using the matrix determinant lemma as below:

$$p_{A+\mathbf{v}\mathbf{v}^T}(x) = \det(xI - A - \mathbf{v}\mathbf{v}^T) = p_A(x) \left(1 - \sum_j \frac{\langle \mathbf{v}, u_j \rangle^2}{x - \lambda_i} \right)$$

where λ_i are the eignevalues of A and u_j are the corresponding eigenvectors. The polynomial $p_{A+\mathbf{v}\mathbf{v}^T}(x)$ has 2 kinds of zeros λ:

1. $p_A(\lambda) = 0$
   These are equal to the eigenvalues λ_j of A for which the added vector $\mathbf{v}$ is orthogonal to the corresponding eigenvector u_j, and do not 'move' upon adding $\mathbf{v}\mathbf{v}^T$.

2. $p_A(\lambda) \neq 0$ and

   $$f(\lambda) = \left(1 - \sum_j \frac{\langle \mathbf{v}, u_j \rangle^2}{\lambda - \lambda_j} \right) = 0.$$

   These are the eigenvalues which have moved and strictly interlace the old eigenvalues.

In the sense of a physical model, we could interpret the eigenvalues λ as charged particles lying on a slope. On the slope are n fixed, chargeless barriers at the initial eigenvalues λ_j and each particle is resting against one of the barriers under the influence of gravity. Adding the vector $\mathbf{v}\mathbf{v}^T$ corresponds to placing a charge of $\langle \mathbf{v}, u_j \rangle^2$ on the barrier corresponding to λ_j. The charges on the barriers repel those on the eigenvalues with a force that is proportional to the charge on the barrier and inversely proportional to the distance from the barrier. The force from barrier j is given by

$$\frac{\langle \mathbf{v}, u_j \rangle^2}{\lambda - \lambda_j},$$

a quantity which is positive for λ_j 'below' λ, which are pushing the particle 'upward', negative otherwise. The eigenvalues moves up the slope until they reach an equilibrium where the reuplsive forces from the barriers cancel the effect of gravity, which we take to be a $+1$ in the downward direction. Thus the equilibrium condition corresponds exactly to having the total 'downward pull' $f(\lambda)$ equal to zero.

With the physical model to visualize, we can consider what happens to the eigenvalues of A when we add a random vector from our set $\{\mathbf{v}_i\}$. The first observation is that for any eigenvector u_j, the expected projection of a randomly chosen vector $\mathbf{v} \in \{\mathbf{v}_i\}_{i \leq m}$ is

$$\mathbb{E}_{\mathbf{v}}\left(\langle \mathbf{v}_i, u_j \rangle^2\right) = \frac{1}{m}\sum_i \langle \mathbf{v}_i, u_j \rangle^2 = \frac{1}{m} u_j^T \left(\sum_i \mathbf{v}_i \mathbf{v}_i^T\right) u_j = \frac{||u_j||^2}{m} = \frac{1}{m}$$

This does not mean that there is any single vector $\mathbf{v}_i$ in our set that realises this expected behaviour of equal projections onto the eigenvectors. But if we were to add such a vector in the physical model, we would add equal charges of $1/m$ to each of the barriers and we would expect all the eigenvalues of A to drift forward 'steadily'. One might expect that after sufficiently many iterations, the eigenvalues would all move forward together, with no eigenvalue too far ahead or behind. We would end up in a position where $\lambda_{\max}/\lambda_{\min}$ is bounded.

This intuition turns out to be right. Adding a vector with equal projections change the characteristic polynomial in the following manner:

$$p_{A+\mathbf{v}_{\text{avg}}\mathbf{v}_{\text{avg}}^T}(x) = p_A(x)\left(1 - \sum_j \frac{1/m}{x - \lambda_j}\right) = p_A(x) - (1/m)p'_A(x),$$

since $p'_A(x) = \sum_j \prod_{i \neq j}(x - \lambda_i)$. If we start with $A = 0$, which has the chacteristic polynomial $p_0(x) = x^n$, then after k iterations, we obtain the polynomial

$$p_k(x) = (I - (1/m)D)^k x^n$$

where D in this case is the derivative with respect to x. Iterating the operator $(I - \alpha D)$ for any $\alpha > 0$ generates a standard family of orthogonal polynomials, the associated Laguerre polynomials [7]. Based on [7], after $k = dn$ iterations, the ratio of the larges to the smallest zero is known to be

$$\frac{d+1+2\sqrt{d}}{d+1-2\sqrt{d}}$$

which is what we desire.

Thus to prove this theorem, we will need to show that we can choose a sequence of vectors that realizes the expected behaviour, as long as we are allowed to add arbitrary fractional amounts of $\mathbf{v}_i\mathbf{v}_i^T$ via the weights $s_i \geq 0$. We will control the eigenvalues of our matrix by maintaining two barriers as in the physical model. Thus we will also introduce barrier functions.

1.5.2 Barrier Functions

We begin by defining two barrier potential functions that measure the quality of the eigenvalues of a matrix.

Definition 1.6. *For $u, l \in \mathbb{R}$ and A a symmetric matrix with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, we define*

$$\Phi^u(A) \overset{def}{=} \mathrm{Tr}(uI - A)^{-1} = \sum_i \frac{1}{u - \lambda_i} \quad \textit{(Upper potential)}$$

$$\Phi_l(A) \overset{def}{=} \mathrm{Tr}(A - lI)^{-1} = \sum_i \frac{1}{\lambda_i - l} \quad \textit{(Lower potential)}$$

As long as $A \prec uI$ and $A \succ lI$, these potential functions would measure how far the eigenvalues of A are from the barriers u and l. Particularly, they blow up as any eigenvalue approaches a barrier, since $uI - A$ or $A - lI$ approaches a singular matrix. Their strength lies in that they could reflect the location of all eigenvalues simultaneously. For example, $\Phi^u(A) \leq 1$ implies that no λ_i is within distance of one of u, no 2 λ_is are at distance 2, and no k are at distance k. In terms of physical model, the upper potential is equal to the total repulsion of the eigenvalues of A from upper barrier u, while the lower potential is the analogous quantity.

To prove the theorem, we will build the sum $\sum_i s_i\mathbf{v}_i\mathbf{v}_i^T$ iteratively, adding one vector at a time. Particularly, we will construct a sequence of matrices

$$0 = A^{(0)}, A^{(1)}, ..., A^{(Q)}$$

along with positive constants $u_0, l_0, \delta_U, \delta_L, \epsilon_U, \epsilon_L$ which satisfy the following conditions:

1. Initially, the barriers are at $u = u_0$ and $l = l_0$ and the potentials are at

$$\Phi^{u_0}(A^{(0)}) = \epsilon_U, \quad \Phi_{l_0}(A^{(0)}) = \epsilon_L$$

2. Each matrix is obtained by a rank-one update of the previous one.

$$A^{(q+1)} = A^{(q)} + t\mathbf{v}\mathbf{v}^T \quad \exists \mathbf{v} \in \{\mathbf{v}_i\}, t \geq 0.$$

3. If we increment the barriers u and l by δ_U and δ_L respectively at each step, then the upper and lower potentials do not increase. For every $q = 0, 1, ..., Q$,

$$\Phi^{u+\delta_U}(A^{(q+1)}) \leq \Phi^u(A^{(q)}) \leq \epsilon_U \quad \text{for } u = u_0 + q\delta_U.$$
$$\Phi_{l+\delta_L}(A^{(q+1)}) \leq \Phi_l(A^{(q)}) \leq \epsilon_L \quad \text{for } l = l_0 + q\delta_L.$$

4. No eigenvalue ever jumps across a barrier. For every $q = 0, 1, ..., Q$,

$$\lambda_{\max}(A^{(q)}) < u_0 + q\delta_U, \quad \lambda_{\min}(A^{(q)}) > l_0 + q\delta_L$$

Now we just need to choose $u_0, l_0, \delta_U, \delta_L, \epsilon_U, \epsilon_L$ such that after $Q = dn$ steps, the condition number of $A^{(Q)}$ is bounded by

$$\frac{\lambda_{\max}(A^{(Q)})}{\lambda_{\min}(A^{(Q)})} \leq \frac{u_0 + dn\delta_U}{l_0 + dn\delta_L} = \frac{d+1+2\sqrt{d}}{d+1-2\sqrt{d}}$$

By construction, $A^{(Q)}$ is a weighted sum of at most dn of the vectors as we desire.

To show that these conditions can be satisfied, we will introduce the following few lemmas.

Lemma 1.7 (Upper Barrier Shift)**.** *Suppose $\lambda_{\max}(A) < u$ and $\boldsymbol{v}$ is any vector. If*

$$\frac{1}{t} \geq \frac{\boldsymbol{v}^T((u+\delta_U)I - A)^{-2}\boldsymbol{v}}{\Phi^u(A) - \Phi^{u+\delta_U}(A)} + \boldsymbol{v}^T((u+\delta_U)I - A)^{-1}\boldsymbol{v} \stackrel{def}{=} U_A(\boldsymbol{v})$$

then

$$\Phi^{u+\delta_U}(A + t\boldsymbol{v}\boldsymbol{v}^T) \leq \Phi^u(A) \text{ and } \lambda_{\max}(A + t\boldsymbol{v}\boldsymbol{v}^T) < u + \delta_U.$$

That is, if we add $t\boldsymbol{v}\boldsymbol{v}^T$ to A and shift the upper barrier by δ_U, then we do not increase the upper potential.

Proof sketch. Let $u' = u + \delta_U$. By Sherman-Morisson Forumula, we can write the updated potential as:

$$\begin{aligned}
\Phi^{u+\delta_U}(A + t\mathbf{v}\mathbf{v}^T) &= \mathrm{Tr}(u'I - A - t\mathbf{v}\mathbf{v}^T)^{-1} \\
&= Tr\left((u'I - A)^{-1} + \frac{t(u'I - A)^{-1}\mathbf{v}\mathbf{v}^T(u'I - A)^{-1}}{1 - t\mathbf{v}^T(u'I - A)^{-1}\mathbf{v}}\right) \\
&= ... \\
&= \Phi^{u+\delta_U}(A) + \frac{t\mathbf{v}^T(u'I - A)^{-2}\mathbf{v}}{1 - t\mathbf{v}^T(u'I - A)^{-1}\mathbf{v}} \\
&= \Phi^u(A) - (\Phi^u(A) - \Phi^{u+\delta_U}(A)) + \frac{\mathbf{v}^T(u'I - A)^{-2}\mathbf{v}}{1/t - \mathbf{v}^T(u'I - A)^{-1}\mathbf{v}}
\end{aligned}$$

Since $U_A(\mathbf{v}) > \mathbf{v}^T(u'I - A)^{-1}\mathbf{v}$, the last term is finite for $1/t \geq U_A(\mathbf{v})$. By substituting any $1/t \geq U_A(\mathbf{v})$, we find $\Phi^{u+\delta_U}(A + t\mathbf{v}\mathbf{v}^T) \leq \Phi^u(A)$. This also tells us that $\lambda_{\max}(A + t\mathbf{v}\mathbf{v}^T < u + \delta_U$ since if this is not the case, there would be some positive $t' \leq t$ for which $\lambda_{\max}(A + t'\mathbf{v}\mathbf{v}^T) = u + \delta_U$. But at such t', $\Phi^{u+\delta_U}(A + t'\mathbf{v}\mathbf{v}^T)$ would blow up, establishing that it is finite. □

Next, we introduce the analogous version relating to lower barrier.

Lemma 1.8 (Lower Barrier Shift)**.** *Suppose $\lambda_{\min}(A) > l, \Phi_l(A) \leq 1/\delta_L$, and $\boldsymbol{v}$ is any vector. If*

$$0 < \frac{1}{t} \leq \frac{\boldsymbol{v}^T(A - (l + \delta_L)I)^{-2}\boldsymbol{v}}{\Phi_{l+\delta_L}(A) - \Phi_l(A)} - \boldsymbol{v}^T(A - (l + \delta_L)I)^{-1}\boldsymbol{v} \stackrel{def}{=} L_A(\boldsymbol{v})$$

then

$$\Phi_{l+\delta_L}(A + t\boldsymbol{v}\boldsymbol{v}^T) \leq \Phi_l(A) \text{ and } \lambda_{\min}(A + t\boldsymbol{v}\boldsymbol{v}^T) > l + \delta_L$$

That is, if we add $t\boldsymbol{v}\boldsymbol{v}^T$ to A and shift the lower barrier by δ_L, then we do not increase the lower potential.

Proof sketch. Observe $\lambda_{\min}(A) > l$ and $\Phi_l(A) \leq 1/\delta_L$ imp;y that $\lambda_{\min}(A) > l + \delta_L$. So for every $t > 0$, $\lambda_{\min}(A + t\mathbf{v}\mathbf{v}^T) > l + \delta_L$.

Proceeding similarly to the proof for upper potential, let $l' = l + \delta_L$. By Sherman-Morrison, we have

$$\begin{aligned}\Phi_{l+\delta_L}(A + t\mathbf{v}\mathbf{v}^T) &= \mathrm{Tr}(A + t\mathbf{v}\mathbf{v}^T - l'I)^{-1} \\ &= \ldots \\ &= \Phi_l(A) + (\Phi_{l+\delta_L}(A) - \Phi_l(A)) - \frac{\mathbf{v}^T(A - l'I)^{-2}\mathbf{v}}{1/t + \mathbf{v}^T(A - l'I)^{-1}\mathbf{v}}\end{aligned}$$

Rearranging will show that $\Phi_{l+\delta}(A + t\mathbf{v}\mathbf{v}^T) \leq \Phi_l(A)$ when $1/t \leq L_A(\mathbf{v})$. □

Now, the next lemma would identify the conditions in which we can find a single $t\mathbf{v}\mathbf{v}^T$ which allows us to maintain both potentials. However, to prove this, we will first need the following lemma.

Lemma 1.9. *If $\lambda_i > 1$ for all i, $0 \leq \sum_i(\lambda_i - l)^{-1} \leq \epsilon_L$ and $1/\delta_L - \epsilon_L \geq 0$, then*

$$\frac{\sum_i(\lambda_i - l - \delta_L)^{-2}}{\sum_i(\lambda_i - l - \delta_L)^{-1} - \sum_i(\lambda_i - l)^{-1}} - \sum_i \frac{1}{\lambda_i - l - \delta_L} \geq \frac{1}{\delta_L} - \sum_i \frac{1}{\lambda_i - l}$$

Proof. We have

$$\delta_L \leq 1/\epsilon_L \leq \lambda_i - l$$

for every i. So the denominator of the left-most term on the left-hand side is positive and the implied inequality is equivalent to

$$
\begin{aligned}
&\sum_i (\lambda_i - l - \delta_L)^{-2} \\
\geq &\left(\sum_i \frac{1}{\lambda_i - l - \delta_L} - \sum_i \frac{1}{\lambda_i - l} \right) \left(\frac{1}{\delta_L} + \sum_i \frac{1}{\lambda_i - l - \delta_L} - \sum_i \frac{1}{\lambda_i - l} \right) \\
= &\left(\delta_L \sum_i \frac{1}{(\lambda_i - l - \delta_L)(\lambda_i - l)} \right) \left(\frac{1}{\delta_L} + \delta_L \sum_i \frac{1}{(\lambda_i - l - \delta_L)(\delta_i - l)} \right) \\
= &\sum_i \frac{1}{(\lambda_i - l - \delta_L)(\lambda_i - l)} + \left(\delta_L \sum_i \frac{1}{(\lambda_i - l - \delta_L)(\lambda_i - l)} \right)^2
\end{aligned}
$$

Hence,

$$
\delta_L \sum_i \frac{1}{(\lambda_i - l - \delta_L)^2(\lambda_i - l)} \geq \left(\delta_L \sum_i \frac{1}{(\lambda_i - l - \delta_L)(\lambda_i - l)} \right)^2
$$

By Cauchy-Schawrtz inequality, we have

$$
\begin{aligned}
&\left(\delta_L \sum_i \frac{1}{(\lambda_i - l - \delta_L)(\lambda_i - l)} \right)^2 \\
\leq &\left(\delta_L \sum_i \frac{1}{\lambda_i - l} \right) \left(\delta_L \sum_i \frac{1}{(\lambda_i - l - \delta_L)^2(\lambda_i - l)} \right) \\
\leq &(\delta_L \epsilon_L) \left(\delta_L \sum_i \frac{1}{(\lambda_i - l - \delta_L)^2(\lambda_i - l)} \right) \quad (\because \sum_i (\lambda_i - l)^{-1} \leq \epsilon_L) \\
\leq &\delta_L \sum_i \frac{1}{(\lambda_i - l - \delta_L)^2(\lambda_i - l)} \quad (\because \frac{1}{\delta_L} - \epsilon_L \geq 0)
\end{aligned}
$$

□

Now, we are ready to introduce and prove the lemma relating to both barrier shifts.

Lemma 1.10 (Both Barriers)**.** *If* $\lambda_{\max}(A) < u, \lambda_{\min}(A) > l, \Phi^u(A) \leq \epsilon_U, \Phi_l(A) \leq \epsilon_L$ *and* $\epsilon_U, \epsilon_L, \delta_U, \delta_L$ *satisfy*

$$
0 \leq \frac{1}{\delta_U} + \epsilon_U \leq \frac{1}{\delta_L} - \epsilon_L
$$

then there exists an i *and positive* t *for which*

$$
L_A(\boldsymbol{v}_i) \geq 1/t \geq U_A(\boldsymbol{v}_i), \quad \lambda_{\max}(A + t\boldsymbol{v}_i\boldsymbol{v}_i^T) < u + \delta_U, \\
\lambda_{\min}(A + t\boldsymbol{v}_i\boldsymbol{v}_i^T) > l + \delta_L
$$

Proof sketch. We need to show that

$$\sum_i L_A(\mathbf{v}_i) \geq \sum_i U_A(\mathbf{v}_i)$$

By using Lemma 1.7 and Lemma 1.9, we can show that

$$\begin{aligned}\sum_i U_A(\mathbf{v}_i) &\leq 1/\delta_U + \Phi^{u+\delta_U}(A)\\ &\leq 1/\delta_U + \Phi^{u}(A)\\ &\leq 1/\delta_U + \epsilon_U\end{aligned}$$

Similarly, using Lemma 1.8, we can find that

$$\sum_i L_A(\mathbf{v}_i) \geq 1/\delta_L - \epsilon_L.$$

Putting together, we have that

$$\sum_i U_A(\mathbf{v}_i) \leq \frac{1}{\delta_U} + \epsilon_U \leq \frac{1}{\delta_L} - \epsilon_L \leq \sum_i L_A(\mathbf{v}_i)$$

□

With the lemmas established, we are ready to prove Theorem 1.5

Proof of Theorem 1.5. Now we just need to set $\epsilon_U, \epsilon_L, \delta_U, \delta_L$ such that it satisfies Lemma 1.10 and gives a good bound on the condition number. Then we can take $A^{(0)} = 0$ and construct $A^{(q+1)}$ from $A^{(q)}$ by choosing any vector $\mathbf{v}_i$ with

$$L_{A^{(q)}}(\mathbf{v}_i) \geq U_{A^{(q)}}(\mathbf{v}_i)$$

The existence of this vector is guaranteed by Lemma 1.10. Setting $A^{(q+1)} = A^{(q)} + t\mathbf{v}_i\mathbf{v}_i^T$ for any $t \geq 0$ satisfying

$$L_{A^{(q)}}(\mathbf{v}_i) \geq \frac{1}{t} \geq U_{A^{(q)}}(\mathbf{v}_i)$$

It is sufficient to take

$$\delta_L = 1 \qquad \epsilon_L = \frac{1}{\sqrt{d}} \qquad l_0 = -\frac{n}{\epsilon_L}$$

$$\delta_U = \frac{\sqrt{d}+1}{\sqrt{d}-1} \qquad \epsilon_U = \frac{\sqrt{d}-1}{d+\sqrt{d}} \qquad u_0 = \frac{n}{\epsilon_U}$$

We can check that

$$\frac{1}{\delta_U} + \epsilon_U = \frac{\sqrt{d}-1}{\sqrt{d}+1} + \frac{\sqrt{d}-1}{\sqrt{d}(\sqrt{d}+1)} = 1 - \frac{1}{\sqrt{d}} = \frac{1}{\delta_L} - \epsilon_L$$

so that the conditions of Lemma 1.10 is satisfied.

The initial potentials are $\Phi^{\frac{n}{\epsilon_U}}(0) = \epsilon_U$ and $\Phi_{\frac{n}{\epsilon_L}}(0) = \epsilon_L$. After dn steps, we have

$$\begin{aligned}\frac{\lambda_{\max}(A^{(dn)})}{\lambda_{\min}(A^{(dn)})} &\leq \frac{n/\epsilon_U + dn\delta_U}{-n/\epsilon_L + dn\delta_L} \\ &= \frac{\frac{d+\sqrt{d}}{\sqrt{d}-1} + d\frac{\sqrt{d}+1}{\sqrt{d}-1}}{d-\sqrt{d}} \\ &= \frac{d+2\sqrt{d}+1}{d-2\sqrt{d}+1}\end{aligned}$$

as we desire. □

To turn this into an algorithm, one must first compute the vectors $\mathbf{v}_i$, which is done in $O(n^2m)$ time. On each iteration, we must compute $((u+\delta_U)I - A)^{-1}, ((u+\delta_U)I - A)^{-2}$ and the same matrices for the lower potential function. This is done in $O(n^3)$ time.

Finally, we decide which edge to add in each iteration by computing $U_A(\mathbf{v}_i)$ and $L_A(\mathbf{v}_i)$ for each edge, done in $O(n^2m)$ time. Since we run dn number of times, the total time of the algorithm is $O(dn^3m)$.

1.6 Relation to Kadison-Singer conjecture

The Kadison-Singer conjecture is a well-known conjecture that dates back to 1959. The conjecture is equivalent to the Paving conjecture [1]. The following conjecture is due to Weaver [11].

Conjecture 1.11. *There are universal constants $\epsilon > 0, \delta > 0, r \in \mathbb{N}$ for which the following statement holds. If $\boldsymbol{v}_1, ..., \boldsymbol{v} \in \mathbb{R}^n$ satisfy $||\boldsymbol{v}_i|| \leq \delta$ for all i and*

$$\sum_{i \leq m} \boldsymbol{v}_i \boldsymbol{v}_i^T = I,$$

then there is a partition $X_1, ..., X_r$ of $\{1, ..., m\}$ for which

$$\left|\left|\sum_{i \in X_j} \boldsymbol{v}_i \boldsymbol{v}_i^T\right|\right| \leq 1 - \epsilon$$

for every $j = 1, ..., r$.

If we had a version of Theorem 1.5 where assuming $||\mathbf{v}_i|| \leq \delta$ guaranteed that the scalers s_i were all either 0 or some constant $\beta > 0$, and gave a constant approximation of factor $\kappa < \beta$. Then we would have

$$I \preceq \beta \sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^T \preceq \kappa \cdot I$$

for $S = \{i : s_i \neq 0\}$, yielding a proof of conjecture 1.11 with $r = 2$ and $\epsilon = \min\{1 - \frac{\kappa}{\beta}, \frac{1}{\beta}\}$ since

$$\left\|\sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^T\right\| \leq \frac{\kappa}{\beta} \leq 1 - \epsilon$$

and

$$\left\|\sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^T\right\| = 1 - \lambda_{\min}\left(\sum_{i \in S} \mathbf{v}_i \mathbf{v}_i^T\right) \leq 1 - \frac{1}{\beta} \leq 1 - \epsilon$$

2 The Kadison-Singer Problem

The Kadison-Singer problem states the following:

Problem 2.1 (Kadison-Singer Problem)**.** *Does every pure state on the (abelian) von Neumann algebra $\mathbb{D}$ of bounded diagonal operators on ℓ_2 have a unique extension to a pure state on $B(\ell_2)$, the von Neumann algebra of all bounded operators on ℓ_2?*

A state of a von Neumann algebra $\mathcal{R}$ is a linear functional f on $\mathcal{R}$ for which $f(I) = 1$ and $f(T) \geq 0$ whenever $T \geq 0$. The set of states of $\mathcal{R}$ is a convex subset of the dual space of $\mathcal{R}$ which is compact in the w^*-topology. By the Krein-Milman theorem, this convex set is the closed convex hull of its extreme points. The extremal elements in the space of states are called the pure states of $\mathcal{R}$ [6].

Weaver's conjecture, as introduced earlier, is a combinatorial form of the Kadison-Singer problem. A more general form is stated below.

Conjecture 2.2. *There exist universal constants $\eta \geq 2$ and $\theta > 0$ so that the following holds. Let $w_1, \ldots, w_m \in \mathbb{C}^d$ satisfy $\|w_i\| \leq 1$ for all i and suppose*

$$\sum_{i=1}^{m} |\langle u, w_i \rangle|^2 = \eta$$

for every unit vector $u \in \mathbb{C}^d$. Then there exists a partition S_1, S_2 of $\{1, \ldots, m\}$ so that

$$\sum_{i \in S_j} |\langle u, w_i \rangle|^2 \leq \eta - \theta$$

for every unit vector $u \in \mathbb{C}^d$ and each $j \in \{1, 2\}$.

Another equivalent conjecture is the Anderson's paving conjecture, stated below:

Conjecture 2.3. *For every $\epsilon > 0$ there is an $r \in \mathbb{N}$ such that for every $n \times n$ self-adjoint complex matrix T with zero diagonal, there are diagonal projections $P_1, \ldots, P_r$ with $\sum_{i=1}^{r} P_i = I$ such that*

$$\|P_i T P_i\| \leq \epsilon \|T\| \quad \text{for } i = 1, \ldots, r.$$

The main motivation of this section is to use the method of interlacing families of polynomials that we have gone through in past chapters to prove these two conjectures. The main results follows in the following theorem:

Theorem 2.4. *If $\epsilon > 0$ and $v_1, ..., v_m$ are independent random vectors in $\mathbb{C}^d$ with finite support such that*

$$\sum_{i=1}^{m} \mathbb{E}\left(v_i v_i^*\right) = I_d$$

and

$$\mathbb{E}\left(\|v_i\|^2\right) \leq \epsilon, \text{ for all } i$$

then

$$\mathbb{P}\left[\left\|\sum_{i=1}^{m} v_i v_i^*\right\| \leq (1+\sqrt{\epsilon})^2\right] > 0$$

This theorem easily implies the following generalization of Conjecture 2.2.

Corollary 2.5. *Let r be a positive integer and let $u_1, ..., u_m \in \mathbb{C}^d$ be vectors such that*

$$\sum_{i=1}^{m} u_i u_i^* = I$$

and $\|u_i\|^2 \leq \delta$ for all i. Then there exists a partition $\{S_1, ...S_r\}$ of $[m]$ such that

$$\left\|\sum_{i \in S_j} u_i u_i^*\right\| \leq \left(\frac{1}{\sqrt{r}} + \sqrt{\delta}\right)^2 \qquad \text{for } j = 1, ..., r$$

Proof. For each $i \in [m]$ and $k \in [r]$, define $w_{i,k} \in \mathbb{C}^{r \times d}$ to be the direct sum of r vectors from $\mathbb{C}^d$, all of which are $0^d \in \mathbb{C}^d$ except for the k^{th} one which is a copy of $\mathbf{u}_i$. That is,

$$w_{i,1} = \begin{pmatrix} \mathbf{u}_i \\ 0^d \\ \vdots \\ 0^d \end{pmatrix}, w_{i,2} = \begin{pmatrix} 0^d \\ \mathbf{u}_i \\ \vdots \\ 0^d \end{pmatrix}, \text{ and so on.}$$

Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be independent random vectors such that $\mathbf{v}_i$ takes the values $\{\sqrt{r} w_{i,k}\}_{k=1}^{r}$ each with probability $1/r$. These vectors satisfy

$$\mathbb{E}\mathbf{v}_i\mathbf{v}_i^* = \begin{pmatrix} \mathbf{u}_i\mathbf{u}_i^* & 0_{d\times d} & \ldots & 0_{d\times d} \\ 0_{d\times d} & \mathbf{u}_i\mathbf{u}_i^* & \ldots & 0_{d\times d} \\ \vdots & & \ddots & \vdots \\ 0_{d\times d} & 0_{d\times d} & \ldots & \mathbf{u}_i\mathbf{u}_i^* \end{pmatrix} \quad \text{and} \quad \|\mathbf{v}_i\|^2 = r\,\|\mathbf{u}_i\|^2 \leq r\delta.$$

So

$$\sum_{i=1}^{m} \mathbb{E}\left(\mathbf{v}_i \mathbf{v}_i^*\right) = I_{rd}$$

and we can apply Theorem 2.4 with $\epsilon = r\delta$ to show that there exists an assignment of each $\mathbf{v}_i$ so that

$$(1+\sqrt{r\delta})^2 \geq \left\| \sum_{i=1}^{m} \mathbf{v}_i \mathbf{v}_i^* \right\| = \left\| \sum_{k=1}^{r} \sum_{i:\mathbf{v}_i = w_{i,k}} \left(\sqrt{r} w_{i,k}\right) \left(\sqrt{r} w_{i,k}\right)^* \right\|$$

Setting $S_k = \{i : \mathbf{v}_i = w_{i,k}\}$, we obtain

$$\left\| \sum_{i \in S_k} \mathbf{u}_i \mathbf{u}_i^* \right\| = \left\| \sum_{i \in S_k} w_{i,k} w_{i,k^*} \right\| \leq \frac{1}{r} \left\| \sum_{k=1}^{r} \sum_{i:\mathbf{v}_i = w_{i,k}} \left(\sqrt{r} w_{i,k}\right) \left(\sqrt{r} w_{i,k}\right)^* \right\| \leq \left(\frac{1}{\sqrt{r}} + \sqrt{\delta} \right)^2.$$

and this is true for all k. □

Similar to the previous section, if we set $r = 2$ and $\delta = 1/18$, it implies Conjecture 2.2 for $\eta = 18$ and $\theta = 2$. This could also imply Conjecture 2.3 with $r = (6/\epsilon)^4$ which we will show later.

2.1 Linear Algebra Properties

Here, we will state some linear algebra properties that we will use. Previously, we have introduced the rank-one update which we will also use.

For a matrix $M \in \mathbb{C}^{d \times d}$ we write the characteristic polynomial of M in a variable x as

$$\chi[M](x) = \det(xI - M)$$

Theorem 2.6. *For an invertible matrix A and another matrix B of the same dimensions,*

$$\partial_t \det(A + tB) = \det(A) \operatorname{Tr}\left(A^{-1}B\right)$$

We also have the following two trace properties.

For any $k \times n$ matrix A and $n \times k$ matrix B,

$$\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$$

Second is

Lemma 2.7. *If A and B are positive semidefinite matrices of the same dimension, then*

$$\operatorname{Tr}(AB) \geq 0.$$

2.2 Mixed Characteristic Polynomial

Theorem 2.8. *Let $v_1, ..., v_m$ be independent random column vectors in $\mathbb{C}^d$ with finite support. For each i, let $A_i = \mathbb{E}(v_i v_i^*)$. Then,*

$$\mathbb{E}\chi\left[\sum_{i=1}^{m} v_i v_i^*\right](x) = \left(\prod_{i=1}^{m} 1 - \partial_{z_i}\right) \det\left(xI + \sum_{i=1}^{m} z_i A_i\right)\Bigg|_{z_1=...=z_m=0}$$

The expected characteristic polynomial of a sum of independent rank one Hermitian matrices is a function of the covariance matrices A_i which we will call the mixed characteristic polynomial of $A_1, ..., A_m$ and denote by $\mu[A_1, ..., A_m](x)$. The proof of this theorem follows from the following lemma which shows that the random rank one updates of determinants corresponds in a natural way to differential operators.

Lemma 2.9. *For every square matrix A and random vector v, we have*

$$\mathbb{E}[\det(A - vv^*)] = (1 - \partial_t)\det(A + t\mathbb{E}[vv^*])|_{t=0}$$

Proof. Assume A is invertible. By Matrix Determinant Lemma, we have

$$\begin{aligned}
\mathbb{E}(\det(A - vv^*)) &= \mathbb{E}\det(A)\left(1 - v^* A^{-1} v\right) \\
&= \mathbb{E}\det(A)\left(1 - \operatorname{Tr}\left(A^{-1} vv^*\right)\right) \\
&= \det(A) - \det(A)\mathbb{E}\operatorname{Tr}\left(A^{-1} vv^*\right) \\
&= \det(A) - \det(A)\operatorname{Tr}\left(A^{-1}\mathbb{E}vv^*\right)
\end{aligned}$$

On the other hand, by Theorem 2.6, we have

$$(1 - \partial_t)\det(A + t\mathbb{E}vv^*) = \det(A + t\mathbb{E}vv^*) - \det(A)\operatorname{Tr}\left(A^{-1}\mathbb{E}vv^*\right)$$

The claim follows by setting $t = 0$. If A is not invertible, we can choose a sequence of invertible matrices that approach A. Since the identity holds for each matrix in the sequence and the two sides are polynomials in the entries of the matrix, a continuity argument implies that the identity must hold for A as well. □

By applying this lemma inductively, we could prove Theorem 2.8. By using Lemma 1.3 in Chapter 3 and the closure properties of real-stable polynomials, it is immediate that the mixed characteristic polynomial is real rooted.

Corollary 2.10. *The mixed characteristic polynomial of positive semidefinite matrices is real rooted.*

Proof. We know that

$$\det\left(xI + \sum_{i=1}^{m} z_i A_i\right)$$

is real stable. From closure property,

$$\left(\prod_{i=1}^{m} 1-\partial_{z_i}\right)\det\left(xI+\sum_{i=1}^{m} z_iA_i\right)$$

is real stable as well. Now, setting all of the z_i preserves real stability. Hence the resulting polynomial is univariate and is real rooted. $\square$

We use the real rootedness of mixed characteristic polynomials to show that every sequence of independent fisnitely supported random vectors $\mathbf{v}_1, ..., \mathbf{v}_m$ defines an interlacing family. Let l_i be the size of the support of the random vector $\mathbf{v}_i$, and let $\mathbf{v}_i$ take the values $w_{i,1}, ..., w_{i,l_i}$ with probabilities $p_{i,1}, ..., p_{i,l_i}$. For $j_1 \in [l_1], ..., j_m \in [l_m]$, define

$$q_{j_1,...,j_m} = \left(\prod_{i=1}^{m} p_{i,j_i}\right)\chi\left[\sum_{i=1}^{m} w_{i,j_i}w_{i,j_i}^{*}\right](x)$$

Theorem 2.11. *The polynomials $q_{j_1,...,j_m}$ form an interlacing family.*

Proof. For $1 \le k \le m$ and $j_1 \in [l_1], ..., j_k \in [l_k]$, define

$$q_{j_1,...,j_k}(x) = \left(\prod_{i=1}^{k} p_{i,j_i}\right)\mathop{\mathbb{E}}_{v_{k+1},...,v_m}\chi\left[\sum_{i=1}^{k} w_{i,j_i}w_{i,j_i}^{*}+\sum_{i=k+1}^{m} v_iv_i^{*}\right](x)$$

Let

$$q_{\emptyset}(x) = \mathop{\mathbb{E}}_{v_1,...,v_m}\chi\left[\sum_{i=1}^{m} v_iv_i^{*}\right](x)$$

We need to prove that for every partial assignment $j_1, \ldots, j_k$, the polynomials

$$\{q_{j_1,...,j_k,t}(x)\}_{t=1,...,l_{k+1}}$$

have a common interlacing. By Lemma 2.12 in Chapter 2, it suffices to prove that for every nonnegative $\lambda_1, ..., \lambda_{l_{k+1}}$ summing to one, the polynomial

$$\sum_{t=1}^{l_{k+1}} \lambda_t q_{j_1,...,j_k,t}(x)$$

is real rooted.

Let u_{k+1} be a random vector that equals $w_{k+1,t}$ with probability λ_t. Then, the above polynomial equals

$$\left(\prod_{i=1}^{k} p_{i,j_i}\right)\mathop{\mathbb{E}}_{u_{k+1},v_{k+2},...,v_m}\chi\left[\sum_{i=1}^{k} w_{i,j_i}w_{i,j_i}^{*}+u_{k+1}u_{k+1}^{*}+\sum_{i=k+2}^{m} v_iv_i^{*}\right](x)$$

which is a multiple of a mixed characteristic polynomial and is thus real rooted by Corollary 2.10. $\square$

We also introduce the following proposition for real stable polynomials which will be used in the next section

Proposition 2.12. *If $p \in \mathbb{R}[z_1, ..., z_m]$ is real stable and $a \in \mathbb{R}$, then $p|_{z_1=a} = p(a, z_2, ..., z_m) \in \mathbb{R}[z_2, ..., z_m]$ is real stable.*

5 The Multivariate Barrier Argument

In this section we will prove an upper bound on the roots of the mixed characteristic polynomial $\mu[A_1, \ldots, A_m](x)$ as a function of the A_i. We will prove this inductively with a barrier function argument. In our case of interest, $\sum_{i=1}^{m} A_i = I$. The theorem is as follows:

Theorem 2.13. *Suppose $A_1, ..., A_m$ are Hermitian positive semidefinite matrices satisfying $\sum_{i=1}^{m} A_i = I$ and $\mathrm{Tr}(A_i) \le \epsilon$ for all i. Then the largest root of $\mu[A_1, ..., A_m](x)$ is at most $(1 + \sqrt{\epsilon})^2$.*

We begin by deriving a slightly different expression for $\mu[A_1, ..., A_m](x)$ that allows us to reason separately about the effect of each A_i on its roots.

Lemma 2.14. *Let $A_1, ..., A_m$ be Hermitian positive semidefinite matrices. If $\sum_i A_i = I$, then*

$$\mu[A_1, ..., A_m](x) = \left(\prod_{i=1}^{m} 1 - \partial_{y_i}\right) \det\left(\sum_{i=1}^{m} y_i A_i\right)\Bigg|_{y_1=...=y_m=x}$$

Proof. For any differentiable function f, we have

$$\partial_{y_i}(f(y_i))|_{y_i=z_i+x} = \partial_{z_i} f(z_i + x)$$

The lemma follows by substituting $y_i = z_i + x$ into the right hand side expression of the lemma, and observing that it produces the expression on the right hand side of Theorem 2.8. □

Writing

$$\mu[A_1, ..., A_m](x) = Q(x, x, ..., x) \quad (1)$$

where $Q(y_1, ..., y_m)$ is the multivariate polynomial on the right hand side of Lemma 2.14. The bound on the roots of Q is defined as follows:

Definition 2.15. *Let $p(z_1, ..., z_m)$ be a multivariate polynomial. We say that $z \in \mathbb{R}^m$ is above the roots of p if*

$$p(z + t) > 0 \quad \textit{for all} \quad t = (t_1, ..., t_m) \in \mathbb{R}^m, t_i \ge 0$$

We will denote the set of points which are above the roots of p by Ab_p. To prove Theorem 2.13, it is s by suffices by (1) to show that $(1+\sqrt{\epsilon})^2 \cdot \mathbf{1} \in Ab_Q$, where $\mathbf{1}$ is the all-ones vector. This is done by an inductive "barrier function" argument. In particular, we will construct Q by iteratively applying operations of the form $(1-\partial_{y_i})$, and we will track the locations of the roots of the polynomials that arise in this process by studying the evolution of the functions defined below.

Definition 2.16. *Given a real stable polynomial p and a point $z=(z_1,...,z_m) \in \mathrm{Ab}_p$, the barrier function of p in direction i at z is*

$$\Phi_p^i(z) = \frac{\partial_{z_i} p(z)}{p(z)} = \partial_{z_i} \log p(z)$$

Equivalently, we may define Φ_p^i by

$$\Phi_p^i(z_1,\ldots,z_m) = \frac{q'_{z,i}(z_i)}{q_{z,i}(z_i)} = \sum_{j=1}^{r} \frac{1}{z_i-\lambda_j}$$

where the univariate restriction

$$q_{z,i}(t) = p(z_1,...,z_{i-1},t,z_{i+1},...,z_m)$$

has roots $\lambda_1,...,\lambda_r$, which are real by Proposition 2.12.

While Φ_p^i are m-variate functions, the properties that we require of them may be deduced by considering their bivariate restrictions. We establish these properties by exploiting the following powerful characterization of bivariate real stable polynomials.

It is stated in the form we want by Borcea and Brändén [5], and is proven using an adaptation of a result of Helton and Vinnikov [8] by Lewis, Parrilo and Ramana [9].

Lemma 2.17. *If $p(z_1,z_2)$ is a bivariate real stable polynomial of degree exactly d, then there exist $d\times d$ positive semidefinite matrices A,B and a Hermitian matrix C such that*

$$p(z_1,z_2) = \pm\det(z_1A+z_2B+C)$$

Now we will introduce some properties of the barrier functions, which is that, above the roots of a polynomial, they are nonincreasing and convex in every coordinate.

Lemma 2.18. *[10] Suppose p is real stable and $z\in\mathrm{Ab}_p$. Then for all $i,j\le m$ and $\delta\ge 0$,*

$$\Phi_p^i(z+\delta e_j) \le \Phi_p^i(z), \text{ and} \qquad \textit{(monotonicity)}$$
$$\Phi_p^i(z+\delta e_j) \le \Phi_p^i(z) + \delta\cdot\partial_{z_j}\Phi_p^i(z+\delta e_j) \qquad \textit{(convexity).}$$

We are interested in finding points that lie in Ab_Q, where Q is generated by applying several operators of the form $1-\partial_{z_i}$ to the polynomial $\det\left(\sum_{i=1}^{m} z_i A_i\right)$. The purpose of the "barrier functions" Φ_p^i is to allow us to reason about the relationship between Ab_p and $\mathrm{Ab}_{p-\partial_{z_i}p}$; in particular, the monotonicity property alone immediately implies the following statement.

Lemma 2.19. *Suppose that p is real stable, that z is above its roots, and that $\Phi_p^i(z) < 1$. Then z is above the roots of $p - \partial_{z_i} p$.*

Proof. Let t be a nonnegative vector. As Φ is nonincreasing in each coordinate, we have $\Phi_p^i(z+ t) < 1$, hence

$$\partial_{z_i} p(z+t) < p(z+t) \implies \left(p - \partial_{z_i} p\right)(x+t) > 0$$

as desired. □

This lemma allows us to prove that a vector is above the roots of $p - \partial_{z_i} p$. However, it is not strong enough for an inductive argument because the barrier functions can increase with each $1 - \partial_{z_i}$ operator that we apply. Thus require a stronger form, shown in the following lemma.

Lemma 2.20. *Suppose that $p\left(z_1, \ldots, z_m\right)$ is real stable, that $z \in \mathrm{Ab}_p$, and that $\delta > 0$ satisfies*

$$\Phi_p^j(z) \leq 1 - \frac{1}{\delta}$$

Then for all i,

$$\Phi^i_{p-\partial_{z_j} p}\left(z + \delta e_j\right) \leq \Phi_p^i(z)$$

Proof. We will write ∂_i instead of ∂_{z_i} to ease typesetting and notation. We begin by computing an expression for $\Phi^i_{p-\partial_j p}$ in terms of Φ_p^j, Φ_p^i, and $\partial_j \Phi_p^i$:

$$\begin{aligned}
\Phi^i_{p-\partial_j p} &= \frac{\partial_i\left(p - \partial_j p\right)}{p - \partial_j p} \\
&= \frac{\partial_i\left(\left(1 - \Phi_p^j\right) p\right)}{\left(1 - \Phi_p^j\right) p} \\
&= \frac{\left(1 - \Phi_p^j\right)\left(\partial_i p\right)}{\left(1 - \Phi_p^j\right) p} + \frac{\left(\partial_i\left(1 - \Phi_p^j\right)\right) p}{\left(1 - \Phi_p^j\right) p} \\
&= \Phi_p^i - \frac{\partial_i \Phi_p^j}{1 - \Phi_p^j}. \\
&= \Phi_p^i - \frac{\partial_j \Phi_p^i}{1 - \Phi_p^j}
\end{aligned}$$

since $\partial_i \Phi_p^j = \partial_j \Phi_p^i$. We want to show that $\Phi^i_{p-\partial_j p}(z+\delta e_j) \leq \Phi_p^i(z)$. By the above identity this is equivalent to

$$-\frac{\partial_j \Phi_p^i(z+\delta e_j)}{1-\Phi_p^j(z+\delta e_j)} \leq \Phi_p^i(z) - \Phi_p^i(z+\delta e_j)$$

By convexity property of Lemma 2.18.

$$\delta \cdot \left(-\partial_j \Phi_p^i(z+\delta e_j)\right) \leq \Phi_p^i(z) - \Phi_p^i(z+\delta e_j)$$

Thus it is sufficient to establish that

$$-\frac{\partial_j \Phi_p^i(z+\delta e_j)}{1-\Phi_p^j(z+\delta e_j)} \leq \delta \cdot \left(-\partial_j \Phi_p^i(z+\delta e_j)\right)$$

From monotonicity of Lemma 2.18. we know that $\left(-\partial_j \Phi_p^i(z+\delta e_j)\right) \geq 0$. So, we can divide both sides of the above inequality by this term to obtain

$$\frac{1}{1-\Phi_p^i(z+\delta e_j)} \leq \delta$$

Applying Lemma 2.18 once more we observe that $\Phi_p^j(z+\delta e_j) \leq \Phi_p^j(z)$, and conclude that the above ineqaulity is implied by

$$\frac{1}{1-\Phi_p^j(z)} \leq \delta$$

which is implied by the assumption of our lemma. □

Now, we have the sufficient tools to prove Theorems 2.13 and 2.4.

Proof of Theorem 2.13. Let

$$P(y_1, ..., y_m) = \det\left(\sum_{i=1}^m y_i A_i\right)$$

Set

$$t = \sqrt{\epsilon} + \epsilon$$

As all of the matrices A_i are positive semidefinite and

$$\det\left(t \sum_i A_i\right) = \det(tI) > 0$$

the vector $t\mathbf{1}$ is above the roots of P. By Theorem 2.6.

$$\Phi_P^i(y_1, ..., y_m) = \frac{\partial_i P(y_1, ..., y_m)}{P(y_1, ..., y_m)} = \mathrm{Tr}\left(\left(\sum_{i=1}^m y_i A_i\right)^{-1} A_i\right)$$

So

$$\Phi_P^i(t\mathbf{1}) = \mathrm{Tr}\left(A_i\right)/t \leq \epsilon/t = \epsilon/(\epsilon + \sqrt{\epsilon})$$

which we define to be ϕ. Set

$$\delta = 1/(1 - \phi) = 1 + \sqrt{\epsilon}.$$

For $k \in [m]$, define

$$P_k\left(y_1, ..., y_m\right) = \left(\prod_{i=1}^{k} 1 - \partial_{y_i}\right) P\left(y_1, ..., y_m\right)$$

Note that $P_m = Q$. Set x^0 to be the all-t vector, and for $k \in [m]$ define x^k to be the vector that is $t+\delta$ in the first k coordinates and t in the rest. By inductively applying Lemmas 2.19 and 2.20, we prove that x^k is above the roots of P_k, and that for all i

$$\Phi_{P_k}^i\left(x^k\right) \leq \phi.$$

It follows that the largest root of

$$\mu\left[A_1, ..., A_m\right](x) = P_m(x, ..., x)$$

is at most

$$t + \delta = 1 + \sqrt{\epsilon} + \sqrt{\epsilon} + \epsilon = (1 + \sqrt{\epsilon})^2.$$

□

Proof of Theorem 2.4. Let $A_i = \mathbb{E}\left(v_i v_i^*\right)$. We have

$$\mathrm{Tr}\left(A_i\right) = \mathbb{E}\left(\mathrm{Tr}\left(v_i v_i^*\right)\right) = \mathbb{E}\left(v_i^* v_i\right) = \mathbb{E}\left(\|v_i\|^2\right) \leq \epsilon,$$

for all i. The expected characteristic polynomial of the $\sum_i v_i v_i^*$ is the mixed characteristic polynomial $\mu\left[A_1, ..., A_m\right](x)$. Theorem 2.13 implies that the largest root of this polynomial is at most $(1 + \sqrt{\epsilon})^2$.

For $i \in [m]$, let l_i be the size of the support of the random vector v_i, and let v_i take the values $w_{i,1}, ..., w_{i,l_i}$ with probabilities $p_{i,1}, ..., p_{i,l_i}$. Theorem 2.11 tells us that the polynomials $q_{j_1,...,j_m}$ are an interlacing family. So, Theorem 2.11 of Chapter 2 implies that there exist $j_1, ..., j_m$ so that the largest root of the characteristic polynomial of

$$\sum_{i=1}^{m} w_{i,j_i} w_{i,j_i}^*$$

is at most $(1 + \sqrt{\epsilon})^2$. □

References

[1] Charles A Akemann and Joel Anderson. *Lyapunov theorems for operator algebras*, volume 458. American Mathematical Soc., 1991.

[2] Maurice S Bartlett. An inverse matrix adjustment arising in discriminant analysis. *The Annals of Mathematical Statistics*, 22(1):107–111, 1951.

[3] Joshua D Batson, Daniel A Spielman, and Nikhil Srivastava. Twice-ramanujan sparsifiers. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 255–262, 2009.

[4] Adi Ben-Israel and Thomas NE Greville. *Generalized inverses: theory and applications*. Springer Science & Business Media, 2006.

[5] Julius Borcea and Petter Brändén. Multivariate pólya–schur classification problems in the weyl algebra. *Proceedings of the London Mathematical Society*, 101(1):73–104, 2010.

[6] Peter G Casazza and Janet Crandell Tremain. The kadison–singer problem in mathematics and engineering. *Proceedings of the National Academy of Sciences*, 103(7):2032–2039, 2006.

[7] Holger Dette and William J Studden. Some new asymptotic properties for the zeros of jacobi, laguerre, and hermite polynomials. *Constructive Approximation*, 11:227–238, 1995.

[8] J William Helton and Victor Vinnikov. Linear matrix inequality representation of sets. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 60(5):654–674, 2007.

[9] Adrian Lewis, Pablo Parrilo, and Motakuri Ramana. The lax conjecture is true. *Proceedings of the American Mathematical Society*, 133(9):2495–2499, 2005.

[10] Terence Tao. Real stable polynomials and the kadison-singer problem. *URL: https://terrytao.wordpress.com/2013/11/04/real-stable-polynomials-and-the-kadison-singer-problem*, 2013.

[11] Nik Weaver. The kadison–singer problem in discrepancy theory. *Discrete mathematics*, 278(1-3):227–239, 2004.