# Proving Kadison-Singer: A Journey Through Real Stability, Interlacing Families, and Barrier Arguments

Austin Stromme
Math 336, Spring 2015
University of Washington

**Abstract**

Herein we prove the Kadison-Singer Conjecture, following the paper [3] of Marcus, Spielman, and Srivastava closely. We highlight their methods and elaborate on their techniques, for example compiling many closure properties of real stable polynomials. We discuss current developments and identify some limitations of their approach.

# Contents

## 1 Background and Main Theorem

They have In 1959 Kadison and Singer wrote "Extensions of Pure States," a paper on operator theory. In it they posed the fundamental conjecture (it is not important that we know what this actually means):

**Conjecture 1.1** (Kadison-Singer). *Does every pure state on the (abelian) von Neumann algebra $\mathbb{D}$ of bounded diagonal operators on $\ell_2$ have a unique extension to a pure state on $B(\ell_2)$, the von Neumann algebra of all bounded operators on $\ell_2$?*

For over 50 years this problem went unsolved. Along the way numerous equivalent conjectures were identified in many different fields. It was finally reduced by Weaver in [3] to

**Conjecture 1.2** ($KS_2$). *There exist universal constants $\eta \geqslant 2$ and $\theta > 0$ so that the following holds: let $w_1, \ldots, w_m \in \mathbb{C}^d$ satisfy $\|w_i\| \leqslant 1$ 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*

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

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

In the summer of 2013 three mathematicians, Adam Marcus, Daniel Spielman, and Nikhil Srivastava were able to prove $KS_2$ and thus positively resolve Kadison-Singer. They proved the following (see section 2 for notation):

**Theorem 1.3.** *If $\epsilon > 0$ and $v_1, \ldots, v_m$ are independent random vectors in $\mathbb{C}^d$ with finite support such that*

$$\sum_{i=1}^{m} \mathbb{E} v_i v_i^* = I_d,$$

*and*

$$\mathbb{E}\|v_i\|^2 \leqslant \epsilon, \forall i,$$

*then*

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

Let's try and gain some intuition about what this is saying. If the following doesn't make sense (i.e references to graph theory) don't worry and skip it. We can transform Conjecture 1.2 into the following, (we omit the derivation here for brevity):

**Theorem 1.4.** *Given column vectors $v_1, \dots, v_m \in \mathbb{R}^d$ such that*

$$\sum_{i=1}^{m} v_i v_i^* = I_d,$$

*and for all $i$, $\|v_i\| \leqslant \epsilon > 0$, then there is a 2-partitioning $S_1, S_2$ of $[m] := \{1, 2, \dots, m\}$ such that for $j = 1, 2$*

$$1/2 - O(\sqrt{\epsilon}) \leqslant \| \sum_{i \in S_j} v_i v_i^* \| \leqslant 1/2 + O(\sqrt{\epsilon}).$$

This is a much more intuitive formulation of Theorem 1.3, and indeed allows us to informally understand what we are showing as something along the lines of: *given some small random vectors that sum up to identity, we can partition them into two parts that both correspond to about half of what you started with.*

In the context of graph theory this means that we can toss out about half of a graph's edges and retain lots of information about the original graph. Indeed, there are apparently generalizations of Theorem 1.4 that yield a more versatile uncertainty principle that tells you about the "distribution" of uncertainty. Essentially, for *anything* that can be encoded as a quadratic form we can cut things into pieces and preserve some information. Again, if this doesn't make sense don't worry, we're just trying to give intuition about this says. Please see [10] for more information.

# 2 The Game Plan, Notation, and Sufficiency of Our Main Theorem

## 2.1 Plan

Our overall goal is to prove Theorem 1.3. As we saw earlier, this is sufficient to get Conjecture 1.2, and thus positively resolve Kadison-Singer. Along the way we will partially develop the two apparatus of interlacing polynomials and real stable polynomials. We will leverage relevant closure properties of real stable polynomials to get a bound on the largest root of the expected characteristic polynomial, which, via interlacing families of polynomials will yield Theorem 1.3.

## 2.2 Notation

We will use $\|x\|$ to denote the usual 2-norm of $x \in \mathbb{C}^n$. We will *always* mean by $x \in \mathbb{C}^d$ that $x$ is a column vector of size $d$ with complex entries. When we write $I_d$ we mean the $d \times d$ identity matrix. If $A$ is an operator, we agree to let $\|A\| := \max_{\|x\|=1} \|Ax\|$. As usual, for $u \in \mathbb{C}^d$, $u^*$ denotes the complex conjugate transpose. Also when we write $S \in \binom{n}{k}$ we mean that $S \subset [n] := \{1, 2, \dots, n\}$ (we realize this is a slight abuse of the usual notation for $[n]$ but we ask the reader to forgive us in the interests of notation) and $|S| = k$. For a matrix $M \in \mathbb{C}^d$ we say the *characteristic polynomial of $M$ in a variable $x$* is

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

Finally we say that that for two matrices $A, B \in \mathbb{C}^{d\times d}$, $A \preceq B$ if $B - A$ is positive semidefinite.

We say $v_1, \ldots, v_m$ are independent random vectors in $\mathbb{C}^d$ with finite support if the following conditions are met: for each $v_i$ there is a collection of $\ell_i \in \mathbb{N}$ vectors $\{w_{i,j}\}_{j=1}^{\ell_i} \subset \mathbb{C}^d$ and real $\{p_{i,j}\}_{j=1}^{\ell_1} \subset \mathbb{R}_{\geqslant 0}$ such that $\sum_j p_{i,j} = 1$. We want to define what we mean by $\mathbb{P}[P]$ and $\mathbb{E}[f]$ where $P$ is some proposition involving $\{v_i\}$ and $f$ is some map from $\mathbb{C}^{m\times d}$ to $\mathbb{C}^k$. So suppose $P$ is some proposition. Say that $S = \{w_{i,j_i}\}_{i=1}^m$ is a *satisfying assignment* for $P$ if $P$ is true when evaluated with $v_i = w_{i,j_i}$. Then we define

$$\mathbb{P}[P] := \sum_{S \text{ a satisfying assignment for } P} \prod_{i=1}^m p_{i,j_i}.$$

Next suppose $f\colon \mathbb{C}^{m\times d} \to \mathbb{C}^k$ is some map into $\mathbb{C}^k$ ($k$ is just some natural number) that takes arguments of the form $(u_1, u_2, \ldots, u_m)$ for $u_j \in \mathbb{C}^d$ for $j = 1, \ldots, m$. Then we define

$$\mathbb{E}[f] := \sum_{j_1=1}^{\ell_1} \sum_{j_2=1}^{\ell_2} \cdots \sum_{j_m=1}^{\ell_m} f(w_{1,j_1}, w_{2,j_2}, \ldots, w_{m,j_m}) \prod_{i=1}^m p_{i,j_i}.$$

Finally if $v_1, \ldots, v_m$ independent random vectors in $\mathbb{C}^d$ and we write $\mathbb{E}_{v_1,\ldots,v_{m-1}}[f]$ or $\mathbb{P}_{v_1,\ldots,v_{m-1}}[P]$ we mean the above except with the random vectors restricted to just the first $m - 1$.

### 2.3 Sufficiency of Theorem 1.3

From Theorem 1.3 we can derive Conjecture 1.2:

*Proof of Conjecture 1.2.* Given $w_1, \ldots, w_m \in \mathbb{C}^d$ as specified, and agreeing to let $0^d \in \mathbb{C}^d$ be the 0 vector in $\mathbb{C}^d$, we let

$$u_{i,1} = \begin{pmatrix} w_i \\ 0^d \end{pmatrix}, \quad u_{i,2} = \begin{pmatrix} 0^d \\ w_i \end{pmatrix}.$$

Then let $v_i$ be either $u_{i,1}/\sqrt{\eta}$ or $u_{i,2}/\sqrt{\eta}$ with probability 1/2. Observe that in this case

$$\sum_{i=1}^m \mathbb{E} v_i v_i^* = \frac{1}{\eta} \begin{pmatrix} \sum_{i=1}^m w_i w_i^* & 0 \\ 0 & \sum_{i=1}^m w_i w_i^* \end{pmatrix} = 2I_{2d}.$$

We get this because for any unit vector $u \in \mathbb{C}^d$ we have

$$\eta = \sum_{i=1}^{m} |\langle u, w_i\rangle|^2 = \sum_{i=1}^{m} \overline{u^* w_i} u^* w_i = \sum_{i=1}^{m} (u^* w_i)(w_i^* u) = u^* \left( \sum_{i=1}^{m} w_i w_i^* \right) u. \tag{1}$$

This implies $\sum w_i w_i^* = \eta I_d$. Also, we get that $\mathbb{E}\|v_i\|^2 \leqslant 2/\eta$. Since the distribution has finite support by Theorem 1.3 this implies that there is some assignment of $v_i$s so that

$$\begin{aligned}(1+\sqrt{2/\eta})^2 \geqslant \left\| \sum_{i=1}^{m} v_i v_i^* \right\| &= \left\| \sum_{j=1}^{2} \sum_{i:\, v_i = u_{i,j}} 2u_{i,j}u_{i,j}^*/\eta \right\| \\ &= \frac{2}{\eta} \left\| \begin{pmatrix} \sum_{i:\, v_i=u_{i,1}} w_i w_i^* & 0 \\ 0 & \sum_{i:\, v_i=u_{i,2}} w_i w_i^* \end{pmatrix} \right\| \\ &\geqslant \frac{2}{\eta} \left\| \sum_{i:\, v_i = u_{i,j}} w_i w_i^* \right\|, \end{aligned} \tag{2}$$

for $j = 1, 2$. Therefore if we let $S_j = \{i \colon v_i = u_{i,j}\}$ for $j = 1, 2$, we get that

$$\left\| \sum_{i \in S_j} w_i w_i^* \right\| \leqslant (\sqrt{\eta/2} + 1)^2.$$

Setting $\theta = \eta - (\sqrt{\eta/2} + 1)^2 + 2 > 0$ gives that for every unit vector $u \in \mathbb{C}^d$, using the same sort of reasoning as in (1)

$$\begin{aligned} \sum_{i \in S_j} |\langle u, w_i\rangle|^2 &= u^* \left( \sum_{i \in S_j} w_i^* w_i \right) u \\ &\leqslant \|u\| \left\| \left( \sum_{i \in S_j} w_i^* w_i \right) u \right\| \\ &\leqslant (\sqrt{\eta/2} + 1)^2 \leqslant \eta - \theta. \end{aligned}$$

Where we used Cauchy-Schwarz in the first inequality and the definition of the operator norm in the second. Thus 1.2 follows from 1.3. □

# 3 Facts about Real Stable and Interlacing Polynomials

Herein we will review some of the known facts about real stable and interlacing polynomials.

### 3.1 Interlacing Families

**Definition 3.1.** We say that a real-rooted polynomial $p(x) = \alpha_0 \prod_{j=1}^{n-1}(x - \alpha_j)$ *interlaces* a real-rooted polynomial $q(x) = \beta_0 \prod_{j=1}^{n}(x - \beta_j)$ if

$$\beta_1 \leqslant \alpha_1 \leqslant \beta_2 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_{n-1} \leqslant \beta_n.$$

We say that $f_1, \ldots, f_k$ have a *common interlacing* if there is a polynomial $g$ so that $g$ interlaces $f_i$ for each $i$.

In [2] the authors proved the following:

**Lemma 3.2.** *Let $f_1, \ldots, f_k$ be polynomials of the same degree that are real-rooted and have positive leading coefficient. Define $f_\emptyset := \sum_1^k f_i$. If $f_1, \ldots, f_k$ have a common interlacing, then there exists an $i$ so that the largest root of $f_i$ is at most the largest of $f_\emptyset$.*

**Definition 3.3.** Suppose $S_1, \ldots, S_m$ are finite sets and that for every assignment $s_1, \ldots, s_m \in S_1 \times \cdots \times S_m$ let $f_{s_1,\ldots,s_m}(x)$ is a real-rooted degree $n$ polynomial with positive leading coefficient. For a partial assignment $s_1, \ldots, s_k \in S_1 \times \cdots \times S_k$ for $k < m$ we agree to let

$$f_{s_1,\ldots,s_k} := \sum_{(s_{k+1},\ldots,s_m) \in S_{k+1} \times \cdots \times S_m} f_{s_1,\ldots,s_m}.$$

We also let

$$f_\emptyset = \sum f_{(s_k)_{k=1}^m}.$$

We say that the polynomials $\{f_{s_1,\ldots,s_m}\}$ are an *interlacing family* if for all $k = 0, \ldots, m-1$ and all $s_1, \ldots, s_k \in S_1 \times \ldots \times S_k$, the polynomials $\{f_{s_1,\ldots,s_k,t}\}_{t \in S_{k+1}}$ have a common interlacing.

In [3] they give the following result:

**Theorem 3.4.** *Let $S_1, \ldots, S_m$ be finite sets and let $\{f_{s_1,\ldots,s_m}\}$ be an interlacing family of polynomials. Then there exists some $s_1, \ldots, s_m \in S_1 \times \ldots \times S_m$ so that the largest root of $f_{s_1,\ldots,s_m}$ is at most the largest root of $f_\emptyset$.*

*Proof.* We know that $\{f_t\}$ for $t \in S_1$ have a common interlacing and their sum is $f_\emptyset$, so by Lemma 3.2 there is some $s_1$ such that $f_{s_1}$ has all of its roots smaller than the large root of $f_\emptyset$. Proceeding inductively, if $f_{s_1,s_2,\ldots,s_k}$ has its largest root smaller than the largest of $f_\emptyset$, we can use Lemma 3.2 to see there is some $s_{k+1}$ such that $f_{s_1,s_2,\ldots,s_{k+1}}$ has it largest root smaller than the largest of $f_{s_1,s_2,\ldots,s_k}$ which is in turn smaller than the largest of $f_\emptyset$. Note we can apply this because by definition

$$f_{s_1,\ldots,s_k} = \sum_{s_{k+1} \in S_{k+1}} f_{s_1,\ldots,s_{k+1}}.$$

By induction we get the result. □

The following result is going to be very important; interestingly it was apparently independently discovered several times:

**Lemma 3.5.** *Let $f_1, \ldots, f_k$ be (univariate) polynomials of the same degree with positive leading coefficients. Then $f_1, \ldots, f_k$ have a common interlacing iff all convex combinations are real-rooted polynomials.*

Unfortunately the proof is rather long and tedious, and the most long and tedious of the directions is the left implication, which is precisely the direction we will use. See Theorem 2.1 in [4].

## 3.2 Real Stability

The class of real-rooted univariate polynomials is extremely useful; sometimes knowing that the polynomial you are working with has real roots is enough to solve your problem. For our purposes, we noted that Lemma 3.5 is going to be critical to our proof, and the key hypothesis is that we have some collection of real-rooted polynomials. Hence it makes sense to try and find closure properties of real stability. However in our situation we are working in higher dimensions, so it is not enough to just work with real-rootedness for univarite polynomials, we need to generalize this notion to multivariate polynomials. It turns out that the following notion of *real stability* is the correct generalization of real-rootedness.

**Definition 3.6.** We say that a multivariate polynomial $p(z_1, \ldots, z_m) \in \mathbb{C}[z_1, \ldots, z_m]$ is *stable* if whenever $\Im(z_i) > 0$ for all $i$, $p(z_1, \ldots, z_m) \neq 0$. We say that $p$ is *real stable* if it is stable and all of its coefficients are real.

The foundation of our proofs of real stability is the following fundamental result, stated as Proposition 2.5 in [5]:

**Lemma 3.7.** *Let $A_j \in \mathbb{C}^{n \times n}$ for $j = 1, \ldots, m$ be positive semidefinite and $B \in \mathbb{C}^{n \times n}$ be Hermitian. Then*

$$f(z_1, \ldots, z_n) = \det\left(\sum_{j=1}^{m} z_j A_j + B\right)$$

*is either real stable or identically zero.*

As a collection of results, we summarize the closure properties discussed in [6]:

**Theorem 3.8.** *Real stability is preserved under the following:*

1. *Symmetrization: if $p(z_1, \ldots, z_n)$ is real stable then so is $p(z_1, z_1, z_3, \ldots, z_n)$.*
2. *Specialization: If $p(z_1, \ldots, z_n)$ is real stable then so is $p(a, z_2, \ldots, z_n)$ for any $a \in \mathbb{R}$.*

3. *External Field: If* $p(z_1, \ldots, z_n)$ *is real stable then so is* $p(w_1 z_1, \ldots, w_n z_n)$ *for any* $w \in \mathbb{R}^n_{>0}$.

4. *Inversion: If* $p(z_1, \ldots, z_n)$ *is real stable and the degree of* $z_i$ *is* $d_i$ *then* $p(1/z_1, \ldots, 1/z_n) \prod_{i=1}^n z_i^{d_u}$ *is real stable.*

5. *Differentiation 1: If* $p(z_1, \ldots, z_n)$ *is real stable, then so is* $\partial p / \partial z_1$.

6. *Differentiation 2: If* $p(z_1, \ldots, z_n)$ *is real stable, then so is* $(1 - \partial_{z_i})p$.

We will make use of 2 and 6 in our proof of Theorem 1.4.

### 3.3 Relevant Linear Algebra Facts

Finally we recall the following facts from Linear Algebra

**Lemma 3.9.** *If* $A \in \mathbb{C}^{n \times n}$ *is invertible and* $u, v \in \mathbb{C}^n$, *then*

$$\det(A + uv^*) = \det(A)(1 + v^* A^{-1} u).$$

**Lemma 3.10.** *For an invertible* $A \in \mathbb{C}^{n \times n}$ *and Hermitian* $B \in \mathbb{C}^{n \times n}$

$$\partial_t \det(A + tB)\big|_{t=0} = \det(A) \operatorname{Tr}(A^{-1} B).$$

*Proof.* By the spectral theorem we can write $B = \sum_{j=1}^n \lambda_j v_j v_j^*$, so that by 3.9

$$\begin{aligned}
\partial_t \det(A + tB)|_{t=0} &= \det(A) \sum_{j=1}^n \lambda_j v_j^* A^{-1} v_j \prod_{k=1, k \neq j}^n (1 + t\lambda_k v_k^* A^{-1} v_k) \Bigg|_{t=0} \\
&= \det(A) \sum_{j=1}^n \operatorname{Tr}(\lambda_j v_j^* A^{-1} v_j) \\
&= \det(A) \operatorname{Tr}\left( A^{-1} \sum_{j=1}^n \lambda_j v_j^* v_j \right) = \det(A) \operatorname{Tr}(A^{-1} B).
\end{aligned}$$

□

We are now ready to begin building up the results to prove 1.3.

# 4 The Mixed Characteristic Polynomial

We will first show the following:

**Theorem 4.1.** *Let $v_1, \ldots, v_m$ be independent random vectors in $\mathbb{C}^d$ with finite support. For each $i$, let $A_i = \mathbb{E} v_i v_i^*$. Then*

$$\mathbb{E}\chi\left[\sum_{i=1}^m v_i v_i^*\right](x) = \left(\prod_{i=1}^m 1 - \partial_{z_i}\right) \det\left(xI + \sum_{i=1}^m z_i A_i\right)\Bigg|_{z_1=\ldots=z_m=0}. \tag{3}$$

We call the polynomial on the right the *mixed characteristic polynomial* of $A_1, \ldots, A_m$ and denote it by $\mu[A_1, \ldots, A_m](x)$. We first show the following Lemma from [3]:

**Lemma 4.2.** *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 Lemma 3.9 we have that

$$\begin{aligned}\mathbb{E}\det(A - vv^*) &= \mathbb{E}\det(A)(1 + v^* A^{-1} v)\\ &= \mathbb{E}\det(A)(1 - \operatorname{Tr}(A^{-1}vv^*))\\ &= \det(A) - \det(A)\mathbb{E}\operatorname{Tr}(A^{-1}vv^*)\\ &= \det(A) - \partial_t \det(A + t\mathbb{E}vv^*)|_{t=0}.\end{aligned}$$

Where in the last step we used Lemma 3.10. If $A$ is not invertible we can approximate it by invertible matrices for which the desired identity holds. Since the determinant is continuous, the result follows for $A$ too. □

*Proof of Theorem 4.1.* (Due to UW CSE's very own James Lee.) We will apply Lemma 4.2 inductively; i.e. that for every matrix $M$ and all $k$

$$\mathbb{E}\det\left(M - \sum_{i=1}^k v_i v_i^*\right) = \left(\prod_{i=1}^k 1 - \partial_{z_i}\right) \det\left(M + \sum_{i=1}^k z_i A_i\right)\Bigg|_{z_1=\cdots=z_k=0}.$$

For $k = 0$ it is trivial. Assume the induction hypothesis holds for $k - 1$:

$$\begin{aligned}\mathbb{E}\det\left(M - \sum_{i=1}^k v_i v_i^*\right) &= \mathbb{E}_{v_1,\ldots,v_{k-1}}\mathbb{E}_{v_k}\det\left(M - \sum_{i=1}^{k-1} v_i v_i^* - v_k v_k^*\right) && \text{independence}\\ &= \mathbb{E}_{v_1,\ldots,v_{k-1}}(1 - \partial_{z_k})\det\left(M - \sum_{i=1}^{k-1} v_i v_i^* + z_k A_k\right)\Bigg|_{z_k=0} && \text{Lemma 4.2}\\ &= (1 - \partial_{z_k})\mathbb{E}_{v_1,\ldots,v_{k-1}}\det\left(M + z_k A_k - \sum_{i=1}^{k-1} v_i v_i^*\right)\Bigg|_{z_k=0} && \text{linearity of } \partial_{z_k}\\ &= \left(\prod_{i=1}^k 1 - \partial_{z_i}\right)\det\left(M + \sum_{i=1}^k z_i A_i\right)\Bigg|_{z_1=\cdots=z_k=0}.\end{aligned}$$

Which is the result. □

**Corollary 4.3.** *The mixed characteristic polynomial of positive semidefinite matrices is real-rooted.*

*Proof.* By Lemma 3.7 we know that

$$p(z,x) := \det\left(xI + \sum_{i=1}^{m} z_i A_i\right)$$

is real-stable. By Theorem 3.8.6, so is $\left(\prod_{i=1}^{m} 1 - \partial_{z_i}\right) p$. By Theorem 3.8.2, so is the specialization to $z = 0$. But we know that the resulting polynomial is univariate so by the definition of real stability it follows it must be real-rooted (since any imaginary roots it may have must come in conjugate pairs which is impossible by definition of real stability). □

Last we will use the real-rootedness of the mixed characteristic polynomials to show that every sequence of independent finitely supported random vectors $v_1, \ldots, v_m$ defines an interlacing family. Let $l_i$ be the size of the support of the random vector $v_i$, and let $v_i$ take the values $w_{i,1}, \ldots, w_{i,l_i}$ with probabilities $p_{i,1}, \ldots, p_{i,l_i}$. For $j_1 \in [l_1], \ldots, j_m \in [l_m]$. Agree to define

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

**Theorem 4.4.** *The polynomials $q_{j_1,\ldots,j_m}$ form an interlacing family.*

*Proof.* For $1 \leqslant k \leqslant m$ and $j_i \in [l_i]$ for $i = 1, \ldots, k$, define

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

Also agree to let

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

We have to show that for every partial assignment $j_1, \ldots, j_k$ the polynomials $\{q_{j_1,\ldots,j_k,t}(x)\}_{t=1,\ldots,l_{k+1}}$ have a common interlacing. By Lemma 3.5 it suffices to show that any convex combination $\sum_{t=1}^{l_{k+1}} \lambda_t q_{j_1,\ldots,j_k,t}(x)$ is real-rooted. But observe that if we let $u_{k+1}$ be the random vector that is $w_{k+1,t}$ with probability $\lambda_t$. Then

$$\sum_{t=1}^{l_{k+1}} \lambda_t q_{j_1,\ldots,j_k,t}(x) = \left(\prod_{i=1}^{k} p_{i,j_i}\right) \mathbb{E}_{u_{k+1},v_{k+2},\ldots,v_m} \chi\left[\sum_{i=1}^{k} w_{i,j_i} w_{i,j_i}^* + u_{k+1} u_{k+1}^* + \sum_{i=k+2}^{m} v_i v_i^*\right](x).$$

But via rank one updates we can write the above as a constant times a mixed characteristic polynomial, and thus Corollary 4.3 gives the result. □

## 5 The Multivariate Barrier Argument

We will first upper bound the largest root of the mixed characteristic polynomial and then (finally) prove Theorem (1.3).

### 5.1 Upper bounding the largest root of the mixed characteristic polynomial

We want to upper bound the roots of the mixed characteristic polynomial $\mu[A_1, \ldots, A_m](x)$ as a function of the $A_i$, when $\sum A_i = I$. Our main theorem is as follows:

**Theorem 5.1.** *Suppose $A_1, \ldots, A_m$ are Hermitian positive semidefinite matrices satisfying $\sum A_i = I$ and $\mathrm{Tr}(A_i) \leqslant \epsilon$ for all $i$. Then the largest root of $\mu[A_1, \ldots, A_m](x)$ is at most $(1 + \sqrt{\epsilon})^2$.*

**Lemma 5.2.** *Let $A_1, \ldots, A_m$ be Hermitian positive semidefinite matrices. If $\sum_i A_i = I$, then*

$$\mu[A_1, \ldots, A_m](x) = \left( \prod_{i=1}^{m} 1 - \partial_{y_i} \right) \det \left( \sum_{i=1}^{m} y_i A_i \right) \Bigg|_{y_1 = \cdots, y_m = x} \tag{4}$$

*Proof.* This is trivial since for any differentiable real-valued function $f$,

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

This gives the RHS of (3) when applied to the right hand side of (4). □

Let's agree to write

$$\mu[A_1, \ldots, A_m](x) = Q(x, x, \ldots, x)$$

where $Q(y_1, \ldots, y_m)$ is the multivariate polynomial on the right hand side of (4).

**Definition 5.3.** Let $p(z_1, \ldots, z_m)$ be a multivariate polynomial. We say that $z \in \mathbb{R}^n$ is *above* the roots of $p$ if

$$p(z + t) > 0 \quad \text{for all} \quad t \in \mathbb{R}^m_{\geqslant 0}.$$

We denote the set of points above the roots of $p$ by $\mathcal{AB}_p$.

We remark that to prove Theorem 5.1 it is sufficient to show that $(1 + \sqrt{\epsilon})^2 \mathbf{1} \in \mathcal{AB}_Q$, where $\mathbf{1}$ is the all-ones vector. This is easy to see from the definition of $Q$ and its relation to $\mu[A_1, \ldots, A_m]$ above. This is our plan. To achieve this we will use an inductive barrier function argument to construct $Q$ by repeatedly applying operators like $(1 - \partial_{y_i})$, tracking the roots of the polynomials as we go along via the barrier function.

**Definition 5.4.** Given a real stable polynomial $p$ and a point $z = (z_1, \dots, z_m) \in \mathcal{AB}_p$ the *barrier function of $p$ in direction $i$ at $z$* is

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

where $q_{z,i}(t) := p(z_1, \dots, z_{i-1}, t, z_{i+1}, \dots, z_m)$ and has real roots $\lambda_j$, $j \in [r]$ via Theorem 3.8.1,5.

We will leverage the following deep result by Borcea and Branden [5]

**Lemma 5.5.** *If $p(z_1, z_2)$ is a bivariate real stable polynomial of degree exactly $d$, then there are $d \times d$ positive semidefinite matrices $A, B$ and a Hermitian matrix $C$ such that*

$$p(z_1, z_2) = \pm \det(z_1 A + z_2 B + C).$$

We use the following properties of barrier functions, but omit their proofs, which can be found both in [3] and a more elementary version in [9].

**Lemma 5.6.** *Suppose $p$ is real stable and $z \in \mathcal{AB}_p$. Then for all $i, j \leqslant m$ and $\delta \geqslant 0$,*

$$\textit{Monotonicity: } \Phi_p^i(z + \delta e_j) \leqslant \Phi_p^i(z), \tag{5}$$

$$\textit{Convexity: } \Phi_p^i(z + \delta e_j) \leqslant \Phi_p^i(z) + \delta \partial_{z_j} \Phi_p^i(z + \delta e_j). \tag{6}$$

*We add that $\partial_{z_j} \Phi_p^i(z + \delta e_j) \leqslant 0$, so (5) is non-trivial.*

We get that

**Lemma 5.7.** *Suppose that $p$ is real stable, that $z \in \mathcal{AB}_p$, and that $\Phi_p^i(z) < 1$. Then $z \in \mathcal{AB}_{p - \partial_{z_i} p}$.*

*Proof.* Choose any $t \in \mathbb{R}^n_{\geqslant 0}$. Since if $z \in \mathcal{AB}_p$ so is $z + t_i$, we can apply (5) to each coordinate of $t$ iteratively to get that, since $p(z+t) > 0$,

$$(1 - \partial_{z_i}) p(z + t) = p(z + t) \left[1 - \Phi_p^i(z + t)\right] \geqslant p(z + t) \left[1 - \phi_p^i(z)\right] > 0.$$

□

We can improve this to

**Lemma 5.8.** *Suppose that $p(z_1, \dots, z_m)$ is real stable, that $z \in \mathcal{AB}_p$, and $\delta > 0$ satisfies*

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

*Then for all $i$*

$$\Phi_{p - \partial_{z_j} p}^i(z + \delta e_j) \leqslant \Phi_p^i(z).$$

*Proof.* For ease of notation lets write $\partial_{z_j}$ as $\partial_j$. Observe that

$$\begin{aligned}\Phi^i_{p-\partial_j p} &= \frac{\partial_i(p-\partial_j p)}{p-\partial j p} \\ &= \frac{\partial_i((1-\Phi^j_p)p)}{(1-\Phi^j_p)p} \\ &= \frac{(1-\Phi^j_p)\partial_i p + (\partial_i(1-\Phi^j_p))p}{(1-\Phi^j_p)p} \\ &= \Phi^i_p - \frac{\partial_i \Phi^j_p}{1-\Phi^j_p} = \Phi^i_p - \frac{\partial_j \Phi^i_p}{1-\Phi^j_p}.\end{aligned}$$

Where in the last equality we used the fact that

$$\partial_i \Phi^j_p = \partial_i \partial_j \ln(p) = \partial_j \partial_i \ln(p) = \partial_j \Phi^i_p.$$

Hence it suffices to show that

$$\Phi^i_p(z+\delta e_j) - \frac{\partial_j \Phi^i_p(z+\delta e_j)}{1-\Phi^j_p(z+\delta e_j)} \leqslant \Phi^i_p(z) \iff -\frac{\partial_j \Phi^i_p(z+\delta e_j)}{1-\Phi^j_p(z+\delta e_j)} \leqslant \Phi^i_p(z) - \Phi^i_p(z+\delta e_j)$$

By the convexity of $\Phi^i_p$, we know that

$$\delta(-\partial_j \Phi^i_p(z+\delta e_j)) \leqslant \Phi^i_p(z) - \Phi^i_p(z+\delta e_j).$$

Hence it suffices to show that

$$-\frac{\partial_j \Phi^i_p(z+\delta e_j)}{1-\Phi^j_p(z+\delta e_j)} \leqslant \delta(-\partial_j \Phi^i_p(z+\delta e_j))$$

But this is equivalent, since by 6, $(-\partial_j \Phi^i_p(z+\delta e_j)) \geqslant 0$, to

$$\frac{1}{1-\Phi^j_p(z+\delta e_j)} \leqslant \delta.$$

Which is yielded by our hypothesis. □

*Proof of Theorem 5.1.* Let

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

And let $t = \sqrt{\epsilon} + \epsilon$. For any $x \in \mathbb{R}^m_{\geqslant 0}$ it is not hard to see that because the $A_i$ are positive semidefinite,

$$P(t+x) = \det\left(\sum_i (t+x_i)A_i\right) = \det\left(tI + \sum_i x_i A_i\right) \geqslant \det(tI) > 0.$$

Thus $t\mathbf{1} \in \mathcal{AB}_P$. By Theorem 3.10 it follows that

$$\Phi_P^i(y_1, \ldots, y_m) = \mathrm{Tr}\left(\left(\sum_{j=1}^m y_j A_j\right)^{-1} A_j\right)$$

Therefore

$$\Phi_P^i(t\mathbf{1}) = \mathrm{Tr}(A_i)/t \leqslant \epsilon/t = \frac{\epsilon}{\epsilon + \sqrt{\epsilon}}.$$

We let this last quantity be $\phi$. Set $\delta = 1/(1-\phi) = 1 + \sqrt{\epsilon}$. For $k \in [m]$ define

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

Observe $P_m = Q$. 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 5.7 and 5.8 we get that $x^k \in \mathcal{AB}_{P_k}$ and that for all $i$ $\Phi_{P_k}^i(x^k) \leqslant \phi$, respectively.

Thus $x^m \in \mathcal{AB}_{P_m} = \mathcal{AB}_Q$, so that the largest root of $\mu[A_1, \ldots, A_m](x) = P_m(x, \ldots, x)$ is at most

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

□

## 5.2 Proof of Main Theorem

*Proof of Theorem 1.3.* Let $A_i$ be $\mathbb{E} v_i v_i^*$. Then

$$\mathrm{Tr}(A_i) = \mathbb{E}\,\mathrm{Tr}(v_i v_i^*) = \mathbb{E} v_i^* v_i = \mathbb{E}\|v_i\|^2 \leqslant \epsilon.$$

The expected characteristic polynomial of $\sum_i v_i v_i^*$ is the mixed characteristic polynomial $\mu[A_1, \ldots, A_m](x)$, by definition. By Theorem 5.1, the largest root of the expected characteristic polynomial is $(1 + \sqrt{\epsilon})^2$. Thus we can use the notation from Theorem 4.4 to see that $q_{j_1, \ldots, j_m}$ are an interlacing family, and thus by Theorem 3.4 we see that there is some $j_1, \ldots, j_m$ such that the largest root of

$$\chi\left[\sum_{i=1}^m w_{i,j_i} w_{i,j_i}^*\right](x)$$

is at most $(1 + \sqrt{\epsilon})^2$. Hence Theorem 1.3, and thus Kadison-Singer. □

## 6 Conclusion

In conclusion we highlight the machinery that we used and discuss its extensions and limitations. We used our barrier function argument to extend real stability to get Theorem 5.1 and then used this to prove 1.3 by using real stability and interlacing families of polynomials. There are whole research programs dedicated to exploring the properties of real stable polynomials and their generalization to hyperbolic polynomials, see especially Petter Branden's webpage. Indeed Branden and co. were able to apply some of his work on real stability to resolve an important conjecture called the Monotone Permanent conjecture, see [9]. Hence apart from just proving Kadison-Singer, we feel that we laid some of the groundwork for some machinery that looks promising with respect to as yet unresolved problems.

For applications of Kadison-Singer to modern problems, please see [7] and the entire lecture series given at the University of Washington Computer Science department in Spring of 2015 within course 599 by Shayan Gharan at `http://homes.cs.washington.edu/~shayan/courses/cse599/index.html`. There are important applications of this work to problems in theoretical computer science and spectral graph theory. Our original intention was to present such an application but it became too lengthy and off-topic to develop all the machinery.

The most unfortunate thing about this proof is that it is non-constructive, which from the computer science point of view is very important. So future directions are twofold: practically it would be very useful to find a construction for the partition of vectors, and less practically it seems that further development and exploitation of the theory of real stability/hyperbolicity (the generalization of real stability) and interlacing families of polynomials could pay off.

## References

[1] R. Kadison, I. Singer. "Extensions of Pure States." *Amer. J. Math.* **81** (1959), 547-564.

[2] A. Marcus, D. Spielman, N. Srivastava. "Interlacing Families I: Bipartite Ramanujan Graphs of All Degrees." *arXiv preprint arXiv:1304.4132*, 2013.

[3] A. Marcus, D. Spielman, N. Srivastava. "Interlacing Families II: Mixed Characteristic Polynomials and The Kadison-Singer Problem." *arXiv preprint arXiv:1306.3969*, 2013.

[4] J. P. Dedieu. Obreschoff's theorem revisited: what convex sets are contained in this set of hyperbolic polynomials? *Journal of Pure and Applied Algebra*, 81(3):269-278, 1992.

[5] J. Borcea and P. Branden. Applications of stable polynomials to mixed determinants: Johnson's conjectures, and symmetrized Fischer products. *Duke Mathematical Journal*, 143(2):205-223, 2008.

[6] S. O. Gharan. Real Stable and Hyperbolic Polynomials, 2015. from http://homes.cs.washington.edu/-shayan/courses/cse599/adv-approx-10.pdf.

[7] S. O. Gharan. Proof of Kadison-Singer Conjecture and the Extensions, 2015. from http://homes.cs.washington.edu/-shayan/courses/cse599/adv-approx-15.pdf.

[8] T. Tao. Real stable polynomials and the Kadison-Singer problem, 2013. from https://terrytao.wordpress.com/2013/11/04/real-stable-polynomials-and-the-kadison-singer-problem/

[9] P. Branden, J. Haglund, M. Visontai, D. G. Wagner. Proof of the monotone column permanent conjecture. *Trends in Mathematics*, 63-78 (2011).

[10] N. Srivastava, TCS+ talk. *Youtube* `https://www.youtube.com/watch?t=723&v=viGM3PfR2TY`