Annals of Mathematics **182** (2015), 327–350
http://dx.doi.org/10.4007/annals.2015.182.1.8

Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem

By Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava

Abstract

We use the method of interlacing polynomials introduced in our previous article to prove two theorems known to imply a positive solution to the Kadison–Singer problem. The first is Weaver's conjecture KS_2, which is known to imply Kadison–Singer via a projection paving conjecture of Akemann and Anderson. The second is a formulation due to Casazza et al. of Anderson's original paving conjecture(s), for which we are able to compute explicit paving bounds. The proof involves an analysis of the largest roots of a family of polynomials that we call the "mixed characteristic polynomials" of a collection of matrices.

1. Introduction

In their 1959 paper, R. Kadison and I. Singer [35] posed the following fundamental question.

Question 1.1 (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 positive answer to Question 1.1 has been shown to be equivalent to a number of conjectures spanning numerous fields, including Anderson's paving conjectures [4], [5], [6], Weaver's discrepancy theoretic KS_r and KS'_r conjectures [49], the Bourgain–Tzafriri Conjecture [20], [25], and the Feichtinger Conjecture and the R_ε Conjecture [23]; and it was known to be implied by Akemann and Anderson's projection paving conjecture [3, Conj. 7.1.3]. Many approaches to these problems have been proposed; and, under slightly stronger

This research was partially supported by NSF grants CCF-0915487 and CCF-1111257, an NSF Mathematical Sciences Postdoctoral Research Fellowship, Grant No. DMS-0902962, a Simons Investigator Award to Daniel Spielman, and a MacArthur Fellowship.

hypothesis, partial solutions have been found by Berman et al. [12], Bourgain and Tzafriri [21], [20], Paulsen [43], Baranov and Dyakonov [9], Lawton [37], Akemann et al. [2], and Popa [45]. For a discussion of the history and a host of other related conjectures, we refer the reader to [24].

We prove these conjectures by proving Weaver's [49] conjecture KS_2 which, as amended by [49, Th. 2], says

CONJECTURE 1.2 (KS_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*

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

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

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

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

Akemann and Anderson's projection paving conjecture [3, Conj. 7.1.3] follows directly from KS_2 (see [49, p. 229]).

We also give a proof of Anderson's original paving conjecture, which says

CONJECTURE 1.3 (Anderson Paving). *For every $\varepsilon > 0$, there is an $r \in \mathbb{N}$ such that for every $n \times n$ Hermitian matrix T with zero diagonal, there are diagonal projections $P_1, \ldots, P_r$ with $\sum_{i=1}^{r} P_i = I$ such that*

$$\|P_i T P_i\| \leq \varepsilon \|T\| \qquad \textit{for } i = 1, \ldots, r.$$

A similar conjecture is made by Bourgain and Tzafriri [20, Conj. 2.8]. One difference between the paving conjecture and KS_2 is that the paving conjecture can be extended to infinite operators $T \in B(\ell_2)$ by an elementary compactness argument [4], which then gives an immediate resolution of Kadison–Singer in a manner described in the original paper [35, Lemma 5]. On the other hand, the reduction from Kadison–Singer to Akemann and Anderson's projection paving conjecture requires nonelementary operator theory.

Our main result follows. Its proof appears at the end of Section 5.

THEOREM 1.4. *If $\varepsilon > 0$ and $v_1, \ldots, v_m$ are independent random vectors in $\mathbb{C}^d$ with finite support such that*

$$\sum_{i=1}^{m} \mathbb{E}\, v_i v_i^* = I_d \tag{3}$$

and

$$\mathbb{E}\,\|v_i\|^2 \le \varepsilon \quad \textit{for all } i, \tag{4}$$

then

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

The above theorem may be compared to the concentration inequalities of Rudelson [46] and Ahlswede and Winter [1], which imply in our setting that $\|\sum_{i=1}^{m} v_i v_i^*\| \le C(\varepsilon)\cdot \log n$ with high probability. Here we are able to control the deviation at the much smaller scale $(1+\sqrt{\varepsilon})^2$, but only with nonzero probability.

Our theorem easily implies the following generalization of Conjecture 1.2.

COROLLARY 1.5. *Let r be a positive integer, and let $u_1, \ldots, u_m \in \mathbb{C}^d$ be vectors such that*

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

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

$$\left\|\sum_{i\in S_j} u_i {u_i}^*\right\| \le \left(\frac{1}{\sqrt{r}} + \sqrt{\delta}\right)^2 \qquad \textit{for } j = 1, \ldots, r. \tag{5}$$

If we set $r = 2$ and $\delta = 1/18$, this implies Conjecture 1.2 for $\eta = 18$ and $\theta = 2$. To see this, set $u_i = w_i/\sqrt{\eta}$. Weaver's condition (1) becomes $\sum_i u_i u_i^* = I$, and $\delta = 1/\eta$. When we multiply back by η, the result (5) becomes (2) with $\eta - \theta = 16$.

Corollary 1.5 also implies Conjecture 1.3 with $r = (6/\varepsilon)^4$; we defer the (slightly more involved) proof to Section 6.

Proof of Corollary 1.5. For each $i \in [m]$ and $k \in [r]$, define $w_{i,k} \in \mathbb{C}^{rd}$ to be the direct sum of r vectors from $\mathbb{C}^d$, all of which are 0^d (the 0-vector in $\mathbb{C}^d$) except for the kth one, which is a copy of u_i. That is,

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

Now let $v_1, \ldots, v_m$ be independent random vectors such that v_i takes the values $\{\sqrt{r} w_{i,k}\}_{k=1}^{r}$ each with probability $1/r$.

These vectors satisfy

$$\mathbb{E}\, v_i v_i^* = \begin{pmatrix} u_i u_i^* & 0_{d\times d} & \dots & 0_{d\times d} \\ 0_{d\times d} & u_i u_i^* & \dots & 0_{d\times d} \\ \vdots & & \ddots & \vdots \\ 0_{d\times d} & 0_{d\times d} & \dots & u_i u_i^*, \end{pmatrix} \quad \text{and} \quad \|v_i\|^2 = r\,\|u_i\|^2 \le r\delta.$$

So,

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

and we can apply Theorem 1.4 with $\varepsilon = r\delta$ to show that there exists an assignment of each v_i so that

$$(1+\sqrt{r\delta})^2 \ge \left\| \sum_{i=1}^{m} v_i v_i^* \right\| = \left\| \sum_{k=1}^{r} \sum_{i:v_i=w_{i,k}} \left(\sqrt{r} w_{i,k}\right)\left(\sqrt{r} w_{i,k}\right)^* \right\|.$$

Setting $S_k = \{i : v_i = w_{i,k}\}$, for all $k \in [r]$, we obtain

$$\begin{aligned} \left\| \sum_{i\in S_k} u_i {u_i}^* \right\| &= \left\| \sum_{i\in S_k} w_{i,k} {w_{i,k}}^* \right\| \\ &\le \frac{1}{r} \left\| \sum_{k=1}^{r} \sum_{i:v_i=w_{i,k}} \left(\sqrt{r} w_{i,k}\right)\left(\sqrt{r} w_{i,k}\right)^* \right\| \le \left(\frac{1}{\sqrt{r}} + \sqrt{\delta} \right)^2. \qquad \square \end{aligned}$$

2. Overview

We prove Theorem 1.4 using the "method of interlacing polynomials" introduced in [41], which we review in Section 3.1. Interlacing families of polynomials have the property that they always contain at least one polynomial whose largest root is at most the largest root of the sum of the polynomials in the family. In Section 4, we prove that the characteristic polynomials of the matrices that arise in Theorem 1.4 form such a family.

This proof requires us to consider the expected characteristic polynomials of certain sums of independent rank-1 positive semidefinite Hermitian matrices. We call such an expected polynomial a *mixed characteristic polynomial*. To prove that the polynomials that arise in our proof are an interlacing family, we show that all mixed characteristic polynomials are real-rooted. Inspired by Borcea and Brändén's proof of Johnson's Conjecture [17], we do this by constructing multivariate real stable polynomials and then applying operators that preserve real stability until we obtain the (univariate) mixed characteristic polynomials.

We then need to bound the largest root of the expected characteristic polynomial. We do this in Section 5 through a multivariate generalization of the barrier function argument of Batson, Spielman, and Srivastava [10]. The original argument essentially considers the behavior of the roots of a real-rooted univariate polynomial $p(x)$ under the operator $1 - \partial_x$. It does this by keeping track of an upper bound on the roots of the polynomial, along with a measure of how far above the roots this upper bound is. We refer to this measure as the "barrier function."

In our multivariate generalization, we consider a vector x to be *above the roots* of a real stable multivariate polynomial $p(x_1, \ldots, x_m)$ if $p(y_1, \ldots, y_m)$ is nonzero for every vector y that is at least as big as x in every coordinate. The value of our multivariate barrier function at x is the vector of the univariate barrier functions obtained by restricting to each coordinate. We then show that we are able to control the values of the barrier function when operators of the form $1 - \partial_{x_i}$ are applied to the polynomial. Our proof is inspired by a method used by Gurvits [32] to prove the van der Waerden Conjecture and a generalization by Bapat [7] of this conjecture to mixed discriminants. Gurvits's proof examines a sequence of polynomials similar to those we construct in our proof, and it amounts to proving a lower bound on the constant term of the mixed characteristic polynomial.

3. **Preliminaries**

For an integer m, we let $[m] = \{1, \ldots, m\}$. When $z_1, \ldots, z_m$ are variables and $S \subseteq [m]$, we define $z^S = \prod_{i \in S} z_i$.

We write ∂_{z_i} to indicate the operator that performs partial differentiation in z_i, $\partial/\partial z_i$. For a multivariate polynomial $p(z_1, \ldots, z_m)$ and a number x, we write $p(z_1, \ldots, z_m)\Big|_{z_1=x}$ to indicate the restricted polynomial in $z_2, \ldots, z_m$ obtained by setting z_1 to x. We let $\mathsf{Im}(z)$ denote the imaginary part of a complex number z,

As usual, we write $\|x\|$ to indicate the Euclidean 2-norm of a vector x. For a matrix M, we indicate the operator norm by $\|M\| = \max_{\|x\|=1} \|Mx\|$. When M is Hermitian positive semidefinite, we recall that this is the largest eigenvalue of M.

We write $\mathbb{P}$ and $\mathbb{E}$ for the probability of an event and for the expectation of a random variable, respectively.

3.1. *Interlacing families.* We now recall the definition of interlacing families of polynomials from [41] and its main consequence. We say that a univariate polynomial is *real-rooted* if all of its coefficients and roots are real.

Definition 3.1. We say that a real-rooted polynomial

$$g(x) = \alpha_0 \prod_{i=1}^{n-1} (x - \alpha_i)$$

interlaces a real-rooted polynomial $f(x) = \beta_0 \prod_{i=1}^{n}(x - \beta_i)$ if

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

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

In [41], we proved the following elementary lemma, which shows the utility of having a common interlacing.

LEMMA 3.2. *Let $f_1, \ldots, f_k$ be polynomials of the same degree that are real-rooted and have positive leading coefficients. Define*

$$f_\varnothing = \sum_{i=1}^{k} f_i.$$

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 root of $f_\varnothing$.

In many cases of interest, we are faced with polynomials that are indexed naturally by a cartesian product, and it is beneficial to apply Lemma 3.2 inductively to subcollections of the polynomials rather than at once. This inspires the following definition from [41].

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

$$f_{s_1,\ldots,s_k} \stackrel{\text{def}}{=} \sum_{s_{k+1} \in S_{k+1}, \ldots, s_m \in S_m} f_{s_1,\ldots,s_k,s_{k+1},\ldots,s_m}$$

as well as

$$f_\varnothing \stackrel{\text{def}}{=} \sum_{s_1 \in S_1, \ldots, s_m \in S_m} f_{s_1,\ldots,s_m}.$$

We say that the polynomials $\{f_{s_1,\ldots,s_m}\}$ form an *interlacing family* if for all $k = 0, \ldots, m-1$ and all $s_1, \ldots, s_k \in S_1 \times \cdots \times S_k$, the polynomials

$$\{f_{s_1,\ldots,s_k,t}\}_{t \in S_{k+1}}$$

have a common interlacing.

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 $s_1, \ldots, s_m \in S_1 \times \cdots \times S_m$ so that the largest root of $f_{s_1,\ldots,s_m}$ is at most the largest root of $f_\varnothing$.*

Proof. From the definition of an interlacing family, we know that the polynomials $\{f_t\}$ for $t \in S_1$ have a common interlacing and that their sum is $f_\varnothing$. Lemma 3.2 therefore guarantees that one of the polynomials f_t has largest root at most the largest root of $f_\varnothing$. We now proceed inductively. For any $s_1, \ldots, s_k$, we know that the polynomials $\{f_{s_1,\ldots,s_k,t}\}$ for $t \in S_{k+1}$ have a common interlacing and that their sum is $f_{s_1,\ldots,s_k}$. So, for some choice of t (say s_{k+1}), the largest root of the polynomial $f_{s_1,\ldots,s_{k+1}}$ is at most the largest root of $f_{s_1,\ldots,s_k}$. □

We will prove that the polynomials $\{f_s\}$ defined in Section 4 form an interlacing family. According to Definition 3.3, this requires establishing the existence of certain common interlacings. There is a systematic way to show that polynomials have common interlacings by proving that their convex combinations are real-rooted. In particular, the following result seems to have been discovered a number of times. It appears as Theorem 2.1 of Dedieu [27], (essentially) as Theorem 2′ of Fell [29], and as (a special case of) Theorem 3.6 of Chudnovsky and Seymour [26].

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 if and only if $\sum_{i=1}^k \lambda_i f_i$ is real-rooted for all values of $\lambda_i \geq 0$ such that $\sum_{i=1}^k \lambda_i = 1$.*

3.2. *Stable polynomials.* Our results employ tools from the theory of stable polynomials, a generalization of real-rootedness to multivariate polynomials. We recall that a polynomial $p(z_1, \ldots, z_m) \in \mathbb{C}[z_1, \ldots, z_m]$ is *stable* if it is identically zero or if whenever $\mathsf{Im}(z_i) > 0$ for all i, $p(z_1, \ldots, z_m) \neq 0$. A polynomial p is *real stable* if it is stable and all of its coefficients are real. A univariate polynomial is real stable if and only if it is real-rooted (as defined at the beginning of Section 3.1).

To prove that the polynomials we construct in this paper are real stable, we begin with an observation of Borcea and Brändén [17, Prop. 2.4].

PROPOSITION 3.6. *If $A_1, \ldots, A_m$ are positive semidefinite Hermitian matrices, then the polynomial*

$$\det\left(\sum_i z_i A_i\right)$$

is real stable.

We will generate new real stable polynomials from the one above by applying operators of the form $(1 - \partial_{z_i})$. One can use general results, such as Theorem 1.3 of [16] or Proposition 2.2 of [39], to prove that these operators preserve real stability. It is also easy to prove it directly using the fact that the analogous operator on univariate polynomials preserves stability of polynomials with complex coefficients. The following well-known fact is a consequence of (for instance) Corollary 18.2a in Marden [42], but we include a short proof for completeness.

PROPOSITION 3.7. *Suppose $q \in \mathbb{C}[z]$ is stable. Then $q(z) - q'(z)$ is also stable.*

Proof. It suffices to handle the case when q and q' have no common roots since any common roots are also roots of $q - q'$ and must lie in the closed lower half plane because q is stable.

Assume without loss of generality that q is monic, and let $\lambda_1, \ldots, \lambda_d \in \mathbb{C}$ be the (distinct) roots of q. Observe that the roots of $q(z) - q'(z)$ are simply the zeros of the rational function

$$\begin{aligned}\frac{q(z) - q'(z)}{q(z)} &= 1 - \frac{q'(z)}{q(z)} = 1 - \sum_{i=1}^{d} \frac{1}{z - \lambda_i} \\ &= 1 - \sum_{i=1}^{d} \frac{1}{z - \lambda_i} \cdot \frac{\overline{z} - \overline{\lambda_i}}{\overline{z} - \overline{\lambda_i}} = 1 - \sum_{i=1}^{d} \frac{\overline{z} - \overline{\lambda_i}}{|z - \lambda_i|^2}.\end{aligned}$$

Rearranging and taking conjugates, we find that any zero z must satisfy

$$z \sum_i \frac{1}{|z - \lambda_i|^2} = 1 + \sum_i \frac{\lambda_i}{|z - \lambda_i|^2},$$

whence $\mathsf{Im}(z) \le 0$ since $\frac{1}{|z-\lambda_i|^2} > 0$ and $\mathsf{Im}(\lambda_i) \le 0$ for all i. □

COROLLARY 3.8. *If $p \in \mathbb{R}[z_1, \ldots, z_m]$ is real stable, then*

$$(1 - \partial_{z_1}) p(z_1, \ldots, z_m).$$

is real stable.

Proof. Let $x_2, \ldots, x_m$ be numbers with positive imaginary part. Then, the univariate polynomial

$$q(z_1) = p(z_1, z_2, \ldots, z_m)\Big|_{z_2 = x_2, \ldots, z_m = x_m}$$

is stable, and Proposition 3.7 tells us that $(1 - \partial_{z_1}) q(z_1)$ is also stable. This implies that $(1 - \partial_{z_1}) p$ has no roots in which all of the variables have positive imaginary part. □

We will also use the fact that real stability is preserved under setting variables to real numbers (see, for instance, [48, Lemma 2.4(d)]). This is a simple consequence of the definition and Hurwitz's theorem.

PROPOSITION 3.9. *If $p \in \mathbb{R}[z_1, \ldots, z_m]$ is real stable and $a \in \mathbb{R}$, then $p|_{z_1 = a} = p(a, z_2, \ldots, z_m) \in \mathbb{R}[z_2, \ldots, z_m]$ is real stable.*

3.3. *Facts from linear algebra.* For a matrix $M \in \mathbb{C}^{d \times d}$, we write the characteristic polynomial of M in a variable x as

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

The following identity is sometimes known as the *matrix determinant lemma* or the *rank*-1 *update formula.*

LEMMA 3.10. *If A is an invertible matrix and u, v are vectors, then*

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

We will utilize Jacobi's formula for the derivative of the determinant of a matrix, which can be derived from Lemma 3.10.

THEOREM 3.11. *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 require two standard facts about traces. The first is that for a $k \times n$ matrix A and an $n \times k$ matrix B,

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

The second is

LEMMA 3.12. *If A and B are positive semidefinite matrices of the same dimension, then*

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

One can prove this by decomposing A and B into sums of rank-1 positive semidefinite matrices, using linearity of the trace, and then the first fact about traces.

4. The mixed characteristic polynomial

THEOREM 4.1. *Let $v_1, \ldots, v_m$ 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*

$$\mathbb{E}\,\chi\left[\sum_{i=1}^m v_i v_i^*\right](x) = \left(\prod_{i=1}^m 1 - \partial_{z_i}\right) \det\left(xI + \sum_{i=1}^m z_i A_i\right)\Big|_{z_1=\cdots=z_m=0}. \tag{6}$$

In particular, the expected characteristic polynomial of a sum of independent rank-1 Hermitian matrices is a function of the covariance matrices A_i. We call this polynomial the *mixed characteristic polynomial* of $A_1, \ldots, A_m$, and we denote it by $\mu[A_1, \ldots, A_m](x)$.

The proof of Theorem 4.1 relies on the following simple identity, which shows that random rank-1 updates of determinants correspond in a natural way to differential operators.

LEMMA 4.2. *For every square matrix A and random vector v, we have*

$$\mathbb{E}\det(A - vv^*) = (1 - \partial_t)\det\left(A + t\,\mathbb{E}\, vv^*\right)\Big|_{t=0}. \tag{7}$$

Proof. First, assume A is invertible. By Lemma 3.10, we have

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

On the other hand, by Theorem 3.11, we have

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

The claim follows by setting $t = 0$.

If A is not invertible, we can choose a sequence of invertible matrices that approach A. Since identity (7) 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. $\square$

We prove Theorem 4.1 by applying this lemma inductively.

Proof of Theorem 4.1. We will show by induction on k that for any matrix M,

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

The base case $k = 0$ is trivial. Assuming the claim holds for $i < k$, we have

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

as desired. $\square$

Remark 4.3. The proof of Theorem 4.1 given here (using induction and Lemma 4.2) was suggested to us by James Lee. The slightly longer proof that appeared in our original manuscript was not inductive; rather, it utilized the Cauchy–Binet formula to show the equality of each coefficient.

Now it is immediate from Proposition 3.6 and Corollary 3.8 that the mixed characteristic polynomial is real-rooted.

Corollary 4.4. *The mixed characteristic polynomial of positive semi-definite matrices is real-rooted.*

Proof. Proposition 3.6 tells us that

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

is real stable. Corollary 3.8 tells us that

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

is real stable as well. Finally, Proposition 3.9 shows that setting all of the z_i to zero preserves real stability. As the resulting polynomial is univariate, it is real-rooted. □

Finally, we use the real-rootedness of mixed characteristic polynomials to show that every sequence of independent finitely supported random vectors $v_1, \dots, v_m$ 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 $w_{i,1}, \dots, w_{i,l_i}$ with probabilities $p_{i,1}, \dots, p_{i,l_i}$. For $j_1 \in [l_1], \dots, j_m \in [l_m]$, define

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

Theorem 4.5. *The polynomials $q_{j_1,\dots,j_m}$ form an interlacing family.*

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

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

Also, let

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

We need to prove that for every partial assignment $j_1, \dots, j_k$ (possibly empty), the polynomials

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

have a common interlacing. By Lemma 3.5, it suffices to prove that for every nonnegative $\lambda_1, \dots, \lambda_{l_{k+1}}$ summing to one, the polynomial

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

is real-rooted. To show this, let u_{k+1} be a random vector that equals $w_{k+1,t}$ with probability λ_t. Then the above polynomial equals

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

which is a multiple of a mixed characteristic polynomial and is thus real-rooted by Corollary 4.4. □

5. **The multivariate barrier argument**

In this section we will prove an upper bound on the roots of the mixed characteristic polynomial $\mu[A_1, \dots, A_m](x)$ as a function of the A_i, in the case of interest $\sum_{i=1}^m A_i = I$. Our main theorem of the section is

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

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

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

$$\mu[A_1, \dots, A_m](x) = \left(\prod_{i=1}^{m} 1 - \partial_{y_i}\right) \det\left(\sum_{i=1}^{m} y_i A_i\right)\Big|_{y_1=\dots=y_m=x}. \tag{8}$$

Proof. For any differentiable function f, we have

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

The lemma then follows by substituting $y_i = z_i + x$ into expression (8) and observing that it produces the expression on the right-hand side of (6). □

Let us write

$$\mu[A_1, \dots, A_m](x) = Q(x, x, \dots, x), \tag{9}$$

where

$$Q(y_1, \dots, y_m) = \left(\prod_{i=1}^{m} 1 - \partial_{y_i}\right) \det\left(\sum_{i=1}^{m} y_i A_i\right)$$

is the multivariate polynomial on the right-hand side of (8). The bound on the roots of $\mu\left[A_1,\ldots,A_m\right](x)$ will follow from a "multivariate upper bound" on the roots of Q, defined as follows.

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

$$p(z+t)>0 \qquad \text{for all} \qquad t=(t_1,\ldots,t_m)\in\mathbb{R}^m, t_i\geq 0,$$

i.e., if p is positive on the nonnegative orthant with origin at z.

We will denote the set of points that are above the roots of p by Ab_p. To prove Theorem 5.1, it is sufficient by (9) to show that $(1+\sqrt{\varepsilon})^2\cdot\mathbf{1}\in\mathsf{Ab}_Q$, where $\mathbf{1}$ is the all-ones vector. We will achieve this 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 5.4. Given a real stable polynomial $p\in\mathbb{R}[z_1,\ldots,z_m]$ and a point $z=(z_1,\ldots,z_m)\in\mathsf{Ab}_p$, 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

$$\Phi_p^i(z_1,\ldots,z_m)=\frac{q'_{z,i}(z_i)}{q_{z,i}(z_i)}=\sum_{j=1}^{r}\frac{1}{z_i-\lambda_j}, \tag{10}$$

where the univariate restriction

$$q_{z,i}(t)=p(z_1,\ldots,z_{i-1},t,z_{i+1},\ldots,z_m) \tag{11}$$

has roots $\lambda_1,\ldots,\lambda_r$, which are real by Proposition 3.9.

Although the 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 [16, Cor. 6.7] and is proved using an adaptation of a result of Helton and Vinnikov [34] by Lewis, Parrilo and Ramana [38].

LEMMA 5.5. *For all bivariate real stable polynomials $p(z_1,z_2)$ of degree exactly d, there exist $d\times d$ symmetric positive semidefinite matrices A,B and a symmetric matrix C such that*

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

Remark 5.6. We can also conclude that for every $z_1,z_2>0$, z_1A+z_2B must be positive definite. If this were not the case, then there would be a nonzero vector that is in the nullspace of both A and B. This would cause the degree of the polynomial to be lower than d.

Remark 5.7. Lemma 5.5 guarantees the existence of real symmetric A, B, and C. It is somewhat easier to obtain a representation with complex Hermitian matrices — this was shown earlier by Dubrovin [28] and more recently by Kummer, Plaumann, and Vinzant [36] — and it is worth noting that such a representation is also sufficient for our application.

The two analytic properties of the barrier functions that we use are that, above the roots of a polynomial, they are nonincreasing and convex in every coordinate.

LEMMA 5.8. *Suppose $p \in \mathbb{R}[z_1, \ldots, z_m]$ is real stable and $z \in \mathsf{Ab}_p$. Then for all $i, j \in [m]$ and $\delta \geq 0$,*

$$\Phi_p^i(z + \delta e_j) \leq \Phi_p^i(z) \qquad (\textit{monotonicity}) \tag{12}$$

and

$$\Phi_p^i(z + \delta e_j) \leq \Phi_p^i(z) + \delta \cdot \partial_{z_j} \Phi_p^i(z + \delta e_j) \qquad (\textit{convexity}). \tag{13}$$

Proof. If $i = j$, then we consider the real-rooted univariate restriction $q_{z,i}(z_i) = \prod_{k=1}^{r}(z_i - \lambda_k)$ defined in (11). Since $z \in \mathsf{Ab}_p$, we know that $z_i > \lambda_k$ for all k. Monotonicity follows immediately by considering each term in (10), and convexity is easily established by computing

$$\partial_{z_i}^2 \left(\frac{1}{z_i - \lambda_k} \right) = \frac{2}{(z_i - \lambda_k)^3} > 0 \qquad \text{as } z_i > \lambda_k.$$

In the case $i \neq j$, we fix all variables other than z_i and z_j and consider the bivariate restriction

$$q_{z,ij}(z_i, z_j) = p(z_1, \ldots, z_m).$$

By Lemma 5.5 there are symmetric positive semidefinite B_i, B_j and a symmetric matrix C such that

$$q_{z,ij}(z_i, z_j) = \pm \det(z_i B_i + z_j B_j + C).$$

Remark 5.6 allows us to conclude that the sign is positive: for sufficiently large t, $t(B_i + B_j) + C$ is positive definite and for $t \geq \max(z_i, z_j)$, we have $q_{z,ij}(t, t) > 0$.

The barrier function in direction i can now be simply expressed as

$$\begin{aligned}
\Phi_p^i(z) &= \frac{\partial_{z_i} \det(z_i B_i + z_j B_j + C)}{\det(z_i B_i + z_j B_j + C)} \\
&= \frac{\det(z_i B_i + z_j B_j + C) \operatorname{Tr}\left((z_i B_i + z_j B_j + C)^{-1} B_i\right)}{\det(z_i B_i + z_j B_j + C)} \qquad \text{by Theorem 3.11} \\
&= \operatorname{Tr}\left((z_i B_i + z_j B_j + C)^{-1} B_i\right).
\end{aligned}$$

Let $M = (z_i B_i + z_j B_j + C)$. As $z \in \mathsf{Ab}_p$ and $B_i + B_j$ is positive definite, we can conclude that M is positive definite: if it were not, there would be a $t \geq 0$

for which $\det(M + t(B_i + B_j)) = 0$. We now write

$$\begin{aligned}\Phi_p^i(z+\delta e_j) &= \mathrm{Tr}\left((M+\delta B_j)^{-1}B_i\right)\\ &= \mathrm{Tr}\left(M^{-1}(I+\delta B_j M^{-1})^{-1}B_i\right)\\ &= \mathrm{Tr}\left((I+\delta B_j M^{-1})^{-1}B_i M^{-1}\right).\end{aligned}$$

For δ sufficiently small, we may expand $(I + \delta B_j M^{-1})^{-1}$ in a power series as

$$I - \delta B_j M^{-1} + \delta^2 (B_j M^{-1})^2 + \sum_{\nu \geq 3} (-\delta B_j M^{-1})^{\nu}.$$

Thus,

$$\partial_{z_j}\Phi_p^i(z) = -\mathrm{Tr}\left(B_j M^{-1} B_i M^{-1}\right).$$

To see that this is nonpositive, and thereby prove (12), observe that both B_j and $M^{-1}B_iM^{-1}$ are positive semidefinite, and recall from Lemma 3.12 that the trace of the product of positive semidefinite matrices is nonnegative. To prove convexity, observe that the second derivative is nonnegative because

$$\partial_{z_j}^2\Phi_p^i(z) = \mathrm{Tr}\left((B_jM^{-1})^2B_iM^{-1}\right) = \mathrm{Tr}\left((B_jM^{-1}B_j)(M^{-1}B_iM^{-1})\right)$$

is also the trace of the product of positive semidefinite matrices.

Inequality (13) is equivalent to convexity in direction e_j and may be obtained by observing that $f(x+\delta) \leq f(x) + \delta f'(x+\delta)$ for any convex differentiable f. □

Remark 5.9. It is worth noting that once the determinantal representation is established, convexity and monotonicity also follow directly from the fact that the function $X \mapsto X^{-1}$ is operator monotone and operator convex [13].

There are other ways of proving Lemma 5.8 that go through more elementary techniques than those used by Helton and Vinnikov [34]. James Renegar has pointed out that it follows from Corollary 4.6 of [11]. Terence Tao [47] has also presented a more elementary proof.

Recall that 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 $\det(\sum_{i=1}^m z_iA_i)$. The purpose of the "barrier functions" Φ_p^i is to allow us to reason about the relationship between Ab_p and $\mathsf{Ab}_{p-\partial_{z_i}p}$; in particular, the monotonicity property alone immediately implies the following statement.

LEMMA 5.10. *Suppose that $p \in \mathbb{R}[z_1, \ldots, z_m]$ is real stable, that $z \in \mathsf{Ab}_p$, and that $\Phi_p^i(z) < 1$. Then $z \in \mathsf{Ab}_{p-\partial_{z_i}p}$.*

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

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

as desired. □

Lemma 5.10 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. To remedy this, we will require the barrier functions to be bounded away from 1, and we will compensate for the effect of each $1 - \partial_{z_j}$ operation by shifting our upper bound away from zero in direction e_j. In particular, by exploiting the convexity properties of the Φ_p^i, we arrive at the following strengthening of Lemma 5.10.

LEMMA 5.11. *Suppose that* $p(z_1, \ldots, z_m)$ *is real stable, that* $z \in \mathsf{Ab}_p$, *and that* $\delta > 0$ *satisfies*

$$\Phi_p^j(z) \leq 1 - \frac{1}{\delta}. \tag{14}$$

Then for all i,

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

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

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

as $\partial_i \Phi_p^j = \partial_j \Phi_p^i$. We would like 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 part (13) of Lemma 5.8,

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

and so it suffices to establish that

$$-\frac{\partial_j \Phi_p^i(z + \delta e_j)}{1 - \Phi_p^j(z + \delta e_j)} \leq \delta \cdot (-\partial_j \Phi_p^i(z + \delta e_j)). \tag{15}$$

From part (12) of Lemma 5.8, we know that $(-\partial_j \Phi_p^i(z+\delta e_j)) \geq 0$; so (15) is implied by

$$\frac{1}{1-\Phi_p^i(z+\delta e_j)} \leq \delta. \tag{16}$$

Applying Lemma 5.8 once more we observe that $\Phi_p^j(z+\delta e_j) \leq \Phi_p^j(z)$, and we conclude that (16) is implied by

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

which is implied by (14). □

We now have the necessary tools to prove the main theorem of this section.

Proof of Theorem 5.1. Let

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

Let $t>0$ be a parameter, to be set later. 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 3.11,

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

So,

$$\Phi_P^i(t\mathbf{1}) = \mathrm{Tr}\,(A_i)\,/t \leq \varepsilon/t,$$

which we define to be ϕ. Set

$$\delta = 1/(1-\phi).$$

For $k \in [m]$, define

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

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 5.10 and 5.11, we prove that x^k is above the roots of P_k and that for all i,

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

It follows that the largest root of

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

is at most

$$t+\delta=t+\frac{1}{1-\varepsilon/t}.$$

This is easily seen to be minimized at $t=\sqrt{\varepsilon}+\varepsilon$, yielding the required bound

$$t+\delta=\sqrt{\varepsilon}+\varepsilon+1+\sqrt{\varepsilon}=(1+\sqrt{\varepsilon})^2. \qquad \square$$

Proof of Theorem 1.4. Let $A_i=\mathbb{E}\, v_i v_i^*$. We have

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

for all i.

The expected characteristic polynomial of the $\sum_i v_i v_i^*$ is the mixed characteristic polynomial $\mu\left[A_1,\ldots,A_m\right](x)$. Theorem 5.1 implies that the largest root of this polynomial is at most $(1+\sqrt{\varepsilon})^2$.

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 $w_{i,1},\ldots,w_{i,l_i}$ with probabilities $p_{i,1},\ldots,p_{i,l_i}$. Theorem 4.5 tells us that the polynomials $q_{j_1,\ldots,j_m}$ are an interlacing family. So, Theorem 3.4 implies that there exist $j_1,\ldots,j_m$ 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 $(1+\sqrt{\varepsilon})^2$. $\square$

6. **The paving conjecture**

The main result of this section is the following quantitative version of Conjecture 1.3. Following [22], we will say that a square matrix T can be (r,ε)-*paved* if there are coordinate projections $P_1,\ldots,P_r$ such that $\sum_{i=1}^{r} P_i=I$ and $\|P_i T P_i\|\leq\varepsilon\|T\|$ for all i.

THEOREM 6.1. *For every $\varepsilon>0$, every zero-diagonal Hermitian matrix T can be (r,ε)-paved with $r=(6/\varepsilon)^4$.*

To prove this theorem, we rely on the following result of Casazza et al., which says that paving arbitrary Hermitian matrices can be reduced to paving certain projection matrices. Its short proof is based on elementary linear algebra.

LEMMA 6.2 (Theorem 3 of [22]). *Suppose there is a function $r:\mathbb{R}_+\to\mathbb{N}$ so that every $2n\times 2n$ projection matrix Q with diagonal entries equal to $1/2$ can be $(r(\varepsilon),\frac{1+\varepsilon}{2})$-paved for all $\varepsilon>0$. Then every $n\times n$ Hermitian zero-diagonal matrix T can be $(r^2(\varepsilon),\varepsilon)$-paved for all $\varepsilon>0$.*

Proof of Theorem 6.1. Let Q be an arbitrary $2n \times 2n$ projection matrix with diagonal entries equal to $1/2$. Then $Q = (u_i^* u_j)_{i,j \in [2n]}$ is the Gram matrix of $2n$ vectors $u_1, \ldots, u_{2n} \in \mathbb{C}^n$ with $\|u_i\|^2 = 1/2 = \delta$ and $\sum_i u_i u_i^* = I_n$. Applying Corollary 1.5 to these vectors for a given r yields a partition $S_1, \ldots, S_r$ of $[2n]$. Letting P_k be the projection onto the indices in S_k, we have for each $k \in [r]$,

$$\|P_k Q P_k\| = \left\|(u_i^* u_j)_{i,j \in S_k}\right\| = \left\|\sum_{i \in S_k} u_i u_i{}^*\right\| \leq \left(\frac{1}{\sqrt{r}} + \frac{1}{\sqrt{2}}\right)^2 < \frac{1}{2} + \frac{3}{\sqrt{r}}. \tag{17}$$

Thus every Q can be $(r, \frac{1+\varepsilon}{2})$-paved for $r = 36/\varepsilon^2$. Applying Lemma 6.2 yields Theorem 6.1. □

It is well known that Theorem 6.1 can be extended to arbitrary matrices T with zero diagonal at the cost of a further quadratic loss in parameters: simply decompose $T = A + iB$ for Hermitian zero-diagonal matrices A, B, and take a product of pavings of A and B.

We have not made any attempt to optimize the dependence of r on ε in Theorem 6.1 and leave this as an open question. It is known [22] that it is not possible to do better than $r = 1/\varepsilon^2$.

7. **Conclusion**

7.1. *Ramanujan graphs, matching polynomials, and optimality.* In [41], we introduced interlacing families of polynomials and used them to prove the existence of infinitely many bipartite Ramanujan graphs of every degree, via a conjecture of Bilu and Linial [14]. These are d-regular graphs whose adjacency matrices have nontrivial eigenvalues bounded by $2\sqrt{d-1}$. The relevant expected characteristic polynomials in our proof turn out to be the matching polynomials introduced by Heilmann and Lieb [33], and a simple combinatorial argument in their original paper [33] shows that the roots of these polynomials are bounded by $2\sqrt{d-1}$, which is exactly what we need.

Matching polynomials are a special case of mixed characteristic polynomials (this follows from an identity of Godsil and Gutman [30]), and it turns out that the arguments in this paper, despite being more general, yield a bound on the largest root that is almost as tight as the combinatorial one. In particular, applying Theorem 1.4 in the setting of [41] immediately implies the existence of infinite families of d-regular bipartite graphs with nontrivial eigenvalues at most $2\sqrt{2d} + o(\sqrt{d})$, which is a factor of about $\sqrt{2}$ off from the correct "Ramanujan" bound. On the other hand, A result of Alon and Boppana [15] implies that sufficiently large d-regular graphs must have nontrivial eigenvalues with absolute value at least $2\sqrt{d-1} - o(1)$. Thus the dependence on ε in Theorem 1.4 cannot be made smaller than $1 + \sqrt{2}\sqrt{\varepsilon} + o(\sqrt{\varepsilon})$. We refer the interested reader to [40, §6] for details.

There is also a second and more direct way in which matching polynomials may be seen as mixed characteristic polynomials. When the matrices $A_1, \ldots, A_d$ are diagonal, $\mu[A_1, \ldots, A_d](x)$ is the matching polynomial of the bipartite graph with d vertices on each side in which the edge (i, j) has weight $A_i(j, j)$. When all the matrices have the same trace and their sum is the identity, the graph is regular and our bound on the largest root of the mixed characteristic polynomial agrees to the first order with that obtained by Heilmann and Lieb [33].

7.2. *Mixed discriminants.* When $m = d$, the constant coefficient of the mixed characteristic polynomial of $A_1, \ldots, A_d$ is the mixed discriminant of $A_1, \ldots, A_d$. The mixed discriminant has many definitions, among them

$$D(A_1, \ldots, A_d) = \left(\prod_{i=1}^{d} \partial_{z_i}\right) \det\left(\sum_i z_i A_i\right).$$

See [31] or [8].

When $k < d$, we define

$$D(A_1, \ldots, A_k) = D(A_1, \ldots, A_k, I, \ldots, I)/(d-k)!,$$

where the identity matrix I is repeated $d - k$ times. For example, $D(A_1)$ is just the trace of A_1. With this notation, we can write

$$\mu[A_1, \ldots, A_m](x) = \sum_{k=0}^{d} x^{d-k}(-1)^k \sum_{S \in \binom{[m]}{k}} D((A_i)_{i \in S}).$$

We conjecture that among the families of matrices $A_1, \ldots, A_m$ with $\sum_i A_i = I$ and $\operatorname{Tr}(A_i) \leq \varepsilon$, the largest root of the mixed characteristic polynomial is maximized when as many of the matrices as possible equal $\varepsilon I/d$, another is a smaller multiple of the identity, and the rest are zero. When all of the matrices have the same trace, d/m, this produces a scaled associated Laguerre polynomial $L_d^{m-d}(mx)$. The bound that we prove on the largest root of the mixed characteristic polynomial agrees asymptotically with the largest root of $L_d^{m-d}(mx)$ as d/m is held constant and d grows. Evidence for our conjecture may be found in the work of Gurvits [31], [32], who proves that when $m = d$, the constant term of the mixed polynomial is minimized when each A_i equals I/d.

Two natural questions arise from our work. The first is whether one can design an efficient algorithm to find the partitions and pavings that are guaranteed to exist by Corollary 1.5. The second is broader. There are many operations that are known to preserve real stability and real-rootedness of polynomials (see [39], [16], [18], [19], [44, 48]). For a technique like the method of interlacing polynomials it would be useful to understand what these operations do to the roots and the coefficients of the polynomials.

Acknowledgements. We thank Gil Kalai for making us aware of the Kadison–Singer conjecture and for pointing out the resemblance between the paving conjecture and the sparsification results of [10]. We thank Joshua Batson for discussions of this connection at the beginning of this research effort. We thank James Lee for suggesting the simple proof of Theorem 4.1. Following the initial release of this paper, we received a number of comments and improvements, and we would like to thank everyone who took the time to give us input.

References

[1] R. Ahlswede and A. Winter, Strong converse for identification via quantum channels, *IEEE Trans. Inform. Theory* **48** (2002), 569–579. MR 1889969. Zbl 1071.94530. http://dx.doi.org/10.1109/18.985947.

[2] C. Akemann, J. Anderson, and B. Tanbay, The Kadison-Singer problem for the direct sum of matrix algebras, *Positivity* **16** (2012), 53–66. MR 2892573. Zbl 1270.46056. http://dx.doi.org/10.1007/s11117-010-0109-1.

[3] C. A. Akemann and J. Anderson, Lyapunov theorems for operator algebras, *Mem. Amer. Math. Soc.* **94** (1991), iv+88. MR 1086563. Zbl 0769.46036. http://dx.doi.org/10.1090/memo/0458.

[4] J. Anderson, Extensions, restrictions, and representations of states on C^*-algebras, *Trans. Amer. Math. Soc.* **249** (1979), 303–329. MR 0525675. Zbl 0408.46049. http://dx.doi.org/10.2307/1998793.

[5] J. Anderson, Extreme points in sets of positive linear maps on $\mathcal{B}(\mathcal{H})$, *J. Funct. Anal.* **31** (1979), 195–217. MR 0525951. Zbl 0422.46049. http://dx.doi.org/10.1016/0022-1236(79)90061-2.

[6] J. Anderson, A conjecture concerning the pure states of $B(H)$ and a related theorem, in *Topics in Modern Operator Theory* (Timişoara/Herculane, 1980), *Operator Theory: Adv. Appl.* **2**, Birkhäuser, Boston, 1981, pp. 27–43. MR 0672813. Zbl 0455.47026.

[7] R. B. Bapat, Mixed discriminants of positive semidefinite matrices, *Linear Algebra Appl.* **126** (1989), 107–124. MR 1040776. Zbl 0696.15007. http://dx.doi.org/10.1016/0024-3795(89)90009-8.

[8] R. B. Bapat and T. E. S. Raghavan, *Nonnegative Matrices and Applications*, *Encyclopedia Math. Appl.* **64**, Cambridge Univ. Press, Cambridge, 1997. MR 1449393. Zbl 0879.15015. http://dx.doi.org/10.1017/CBO9780511529979.

[9] A. Baranov and K. Dyakonov, The Feichtinger conjecture for reproducing kernels in model subspaces, *J. Geom. Anal.* **21** (2011), 276–287. MR 2772073. Zbl 1230.30039. http://dx.doi.org/10.1007/s12220-010-9147-y.

[10] J. Batson, D. A. Spielman, and N. Srivastava, Twice-Ramanujan sparsifiers, *SIAM J. Comput.* **41** (2012), 1704–1721. MR 3029269. Zbl 1260.05092. http://dx.doi.org/10.1137/090772873.

[11] H. H. Bauschke, O. Güler, A. S. Lewis, and H. S. Sendov, Hyperbolic polynomials and convex analysis, *Canad. J. Math.* **53** (2001), 470–488. MR 1827817. Zbl 0974.90015. http://dx.doi.org/10.4153/CJM-2001-020-6.

[12] K. Berman, H. Halpern, V. Kaftal, and G. Weiss, Matrix norm inequalities and the relative Dixmier property, *Integral Equations Operator Theory* **11** (1988), 28–48. MR 0920733. Zbl 0647.47007. http://dx.doi.org/10.1007/BF01236652.

[13] R. Bhatia, *Matrix Analysis*, *Grad. Texts in Math.* no. 169, Springer-Verlag, New York, 1997. MR 1477662. Zbl 0863.15001. http://dx.doi.org/10.1007/978-1-4612-0653-8.

[14] Y. Bilu and N. Linial, Lifts, discrepancy and nearly optimal spectral gap, *Combinatorica* **26** (2006), 495–519. MR 2279667. Zbl 1121.05054. http://dx.doi.org/10.1007/s00493-006-0029-7.

[15] R. B. Boppana, Eigenvalues and graph bisection: an average-case analysis, in *Proc.* 28*th IEEE Symposium on Foundations of Computer Science*, 1987, pp. 280–285. http://dx.doi.org/10.1109/SFCS.1987.22.

[16] J. Borcea and P. Brändén, Multivariate Pólya-Schur classification problems in the Weyl algebra, *Proc. Lond. Math. Soc.* **101** (2010), 73–104. MR 2661242. Zbl 1196.47028. http://dx.doi.org/10.1112/plms/pdp049.

[17] J. Borcea and P. Brändén, Applications of stable polynomials to mixed determinants: Johnson's conjectures, unimodality, and symmetrized Fischer products, *Duke Math. J.* **143** (2008), 205–223. MR 2420507. Zbl 1151.15013. http://dx.doi.org/10.1215/00127094-2008-018.

[18] J. Borcea and P. Brändén, The Lee-Yang and Pólya-Schur programs. I. Linear operators preserving stability, *Invent. Math.* **177** (2009), 541–569. MR 2534100. Zbl 1175.47032. http://dx.doi.org/10.1007/s00222-009-0189-3.

[19] J. Borcea and P. Brändén, The Lee-Yang and Pólya-Schur programs. II. Theory of stable polynomials and applications, *Comm. Pure Appl. Math.* **62** (2009), 1595–1631. MR 2569072. Zbl 1177.47041. http://dx.doi.org/10.1002/cpa.20295.

[20] J. Bourgain and L. Tzafriri, Restricted invertibility of matrices and applications, in *Analysis at Urbana*, *Vol. II* (Urbana, IL, 1986–1987), *London Math. Soc. Lecture Note Ser.* **138**, Cambridge Univ. Press, Cambridge, 1989, pp. 61–107. MR 1009186. Zbl 0698.47018. http://dx.doi.org/10.1017/CBO9781107360204.006.

[21] J. Bourgain and L. Tzafriri, On a problem of Kadison and Singer, *J. Reine Angew. Math.* **420** (1991), 1–43. MR 1124564. Zbl 0729.47028. http://dx.doi.org/10.1515/crll.1991.420.1.

[22] P. Casazza, D. Edidin, D. Kalra, and V. I. Paulsen, Projections and the Kadison-Singer problem, *Oper. Matrices* **1** (2007), 391–408. MR 2344683. Zbl 1132.46037. http://dx.doi.org/10.7153/oam-01-23.

[23] P. G. Casazza, O. Christensen, A. M. Lindner, and R. Vershynin, Frames and the Feichtinger conjecture, *Proc. Amer. Math. Soc.* **133** (2005), 1025–1033. MR 2117203. Zbl 1082.46018. http://dx.doi.org/10.1090/S0002-9939-04-07594-X.

[24] P. G. Casazza, M. Fickus, J. C. Tremain, and E. Weber, The Kadison-Singer problem in mathematics and engineering: a detailed account, in *Operator Theory*, *Operator Algebras*, *and Applications*, *Contemp. Math.* **414**, Amer. Math.

Soc., Providence, RI, 2006, pp. 299–355. MR 2277219. Zbl 1082.46018. http://dx.doi.org/10.1090/conm/414/07820.

[25] P. G. Casazza and J. C. Tremain, The Kadison-Singer problem in mathematics and engineering, *Proc. Natl. Acad. Sci. USA* **103** (2006), 2032–2039. MR 2204073. Zbl 1160.46333. http://dx.doi.org/10.1073/pnas.0507888103.

[26] M. Chudnovsky and P. Seymour, The roots of the independence polynomial of a clawfree graph, *J. Combin. Theory Ser. B* **97** (2007), 350–357. MR 2305888. Zbl 1119.05075. http://dx.doi.org/10.1016/j.jctb.2006.06.001.

[27] J.-P. Dedieu, Obreschkoff's theorem revisited: what convex sets are contained in the set of hyperbolic polynomials?, *J. Pure Appl. Algebra* **81** (1992), 269–278. MR 1179101. Zbl 0772.12002. http://dx.doi.org/10.1016/0022-4049(92)90060-S.

[28] B. A. Dubrovin, Matrix finite-gap operators, in *Current Problems in Mathematics*, *Itogi Nauki i Tekhniki* **23**, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1983, pp. 33–78. MR 0734313.

[29] H. J. Fell, On the zeros of convex combinations of polynomials, *Pacific J. Math.* **89** (1980), 43–50. MR 0596914. Zbl 0416.30005. http://dx.doi.org/10.2140/pjm.1980.89.43.

[30] C. D. Godsil and I. Gutman, On the matching polynomial of a graph, in *Algebraic Methods in Graph Theory*, *Vol. I*, *II* (Szeged, 1978), *Colloq. Math. Soc. János Bolyai* **25**, North-Holland, New York, 1981, pp. 241–249. MR 0642044. Zbl 0476.05060.

[31] L. Gurvits, The Van der Waerden conjecture for mixed discriminants, *Adv. Math.* **200** (2006), 435–454. MR 2200852. Zbl 1093.15011. http://dx.doi.org/10.1016/j.aim.2004.12.002.

[32] L. Gurvits, Van der Waerden/Schrijver-Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all, *Electron. J. Combin.* **15** (2008), Research Paper 66, 26, with a corrigendum. MR 2411443. Zbl 1182.15008. Available at http://www.combinatorics.org/Volume_15/Abstracts/v15i1r66.html.

[33] O. J. Heilmann and E. H. Lieb, Theory of monomer-dimer systems, *Comm. Math. Phys.* **25** (1972), 190–232. MR 0297280. Zbl 0228.05131. http://dx.doi.org/10.1007/BF01877590.

[34] J. W. Helton and V. Vinnikov, Linear matrix inequality representation of sets, *Comm. Pure Appl. Math.* **60** (2007), 654–674. MR 2292953. Zbl 1116.15016. http://dx.doi.org/10.1002/cpa.20155.

[35] R. V. Kadison and I. M. Singer, Extensions of pure states, *Amer. J. Math.* **81** (1959), 383–400. MR 0123922. Zbl 0086.09704. http://dx.doi.org/10.2307/2372748.

[36] M. Kummer, D. Plaumann, and C. Vinzant, Hyperbolic polynomials, interlacers, and sums of squares, *Math. Program.* (2013), 1–23 (English). http://dx.doi.org/10.1007/s10107-013-0736-y.

[37] W. Lawton, Minimal sequences and the Kadison-Singer problem, *Bull. Malays. Math. Sci. Soc.* **33** (2010), 169–176. MR 2666420. Zbl 1209.37011.

[38] A. S. Lewis, P. A. Parrilo, and M. V. Ramana, The Lax conjecture is true, *Proc. Amer. Math. Soc.* **133** (2005), 2495–2499. MR 2146191. Zbl 1073.90029. http://dx.doi.org/10.1090/S0002-9939-05-07752-X.

[39] E. H. Lieb and A. D. Sokal, A general Lee-Yang theorem for one-component and multicomponent ferromagnets, *Comm. Math. Phys.* **80** (1981), 153–179. MR 0623156. http://dx.doi.org/10.1007/BF01213009.

[40] A. Marcus, D. A. Spielman, and N. Srivastava, Ramanujan graphs and the solution of the Kadison–Singer problem, in *Proceedings of the International Congress of Mathematicians*, 2014. arXiv 1408.4421.

[41] A. Marcus, D. A. Spielman, and N. Srivastava, Interlacing families I: Bipartite Ramanujan graphs of all degrees, *Ann. of Math.* **182** (2015), 307–325. http://dx.doi.org/10.4007/2015.182.1.7.

[42] M. Marden, *Geometry of Polynomials*, *Math. Surveys Monogr.* **3**, Amer. Math. Soc., Providence, RI, 1966. MR 0225972. Zbl 0162.37101.

[43] V. I. Paulsen, Syndetic sets, paving and the Feichtinger conjecture, *Proc. Amer. Math. Soc.* **139** (2011), 1115–1120. MR 2745663. Zbl 1220.46008. http://dx.doi.org/10.1090/S0002-9939-2010-10600-7.

[44] R. Pemantle, Hyperbolicity and stable polynomials in combinatorics and probability, in *Current Developments in Mathematics*, 2011, Int. Press, Somerville, MA, 2012, pp. 57–123. MR 3098077.

[45] S. Popa, A II_1 factor approach to the Kadison-Singer problem, *Comm. Math. Phys.* **332** (2014), 379–414. MR 3253706. Zbl 06353809. http://dx.doi.org/10.1007/s00220-014-2055-4.

[46] M. Rudelson, Random vectors in the isotropic position, *J. Funct. Anal.* **164** (1999), 60–72. MR 1694526. Zbl 0929.46021. http://dx.doi.org/10.1006/jfan.1998.3384.

[47] T. Tao, Real stable polynomials and the Kadison-Singer problem, 2013. Available at http://terrytao.wordpress.com/2013/11/04/real-stable-polynomials-and-the-kadison-singer-problem/.

[48] D. G. Wagner, Multivariate stable polynomials: theory and applications, *Bull. Amer. Math. Soc.* **48** (2011), 53–84. MR 2738906. Zbl 1207.32006. http://dx.doi.org/10.1090/S0273-0979-2010-01321-5.

[49] N. Weaver, The Kadison-Singer problem in discrepancy theory, *Discrete Math.* **278** (2004), 227–239. MR 2035401. Zbl 1040.46040. http://dx.doi.org/10.1016/S0012-365X(03)00253-X.

(Received: April 17, 2014)

Yale University, New Haven, CT
E-mail: adam.marcus.academic@gmail.com

Yale University, New Haven, CT
E-mail: spielman@cs.yale.edu

Microsoft Research, Bangalore, India
E-mail: knikct@gmail.com