The Kadison-Singer Problem for Strongly Rayleigh Measures and Applications to Asymmetric TSP

Nima Anari * Shayan Oveis Gharan †

Abstract

Marcus, Spielman, and Srivastava in their seminal work [MSS13b] resolved the Kadison-Singer conjecture by proving that the sum of any set of finitely supported independently distributed random vectors with “small” expected squared norm that are in isotropic position (in expectation) attains a “small” spectral norm with a nonzero probability. Their proof crucially employs real stability of polynomials which is the natural generalization of real-rootedness to multivariate polynomials.

Strongly Rayleigh distributions are families of probability distributions whose generating polynomials are real stable [BBL09]. As independent distributions are just special cases of strongly Rayleigh measures, it is a natural question to see if the main theorem of [MSS13b] can be extended to families of vectors assigned to the elements of a strongly Rayleigh distribution.

In this paper we answer this question affirmatively; we show that for any homogeneous strongly Rayleigh distribution where the marginal probabilities are upper bounded by $\epsilon_1$ and any isotropic set of vectors assigned to the underlying elements whose norms are at most $\sqrt{\epsilon_2}$, there is a set in the support of the distribution such that the spectral norm of the sum of the vectors assigned to the elements of the set is at most $O(\epsilon_1 + \epsilon_2)$. We employ our theorem to provide a sufficient condition for the existence of spectrally thin trees. This, together with a recent work of the authors, provide an improved upper bound on the integrality gap of the natural LP relaxation of the Asymmetric Traveling Salesman Problem.

1 Introduction

Marcus, Spielman and Srivastava [MSS13b] in a breakthrough work proved the Kadison-Singer conjecture [KS59] by proving Weaver’s [Wea04] conjecture KS$_2$ and the Akemann and Anderson’s Paving conjecture [AA91]. The following is their main technical contribution.

Theorem 1.1. If $\epsilon > 0$ and $v_1, \ldots, v_m$ are independent random vectors in $\mathbb{R}^d$ with finite support such that

$$\sum_{i=1}^{m} E v_i v_i^\top = I,$$

and for all $i$,

$$E \|v_i\|^2 \leq \epsilon,$$

*Computer Science Division, UC Berkeley. Email: anari@berkeley.edu.
†Computer Science Division, UC Berkeley. Supported by a Miller fellowship. Email: oveisgharan@berkeley.edu.
then
\[ P \left[ \left\| \sum_{i=1}^{m} v_i v_i^T \right\| \leq (1 + \sqrt{\epsilon})^2 \right] > 0. \]

In this paper, we prove an extension of the above theorem to families of vectors assigned to elements of a not necessarily independent distribution.

Let \( \mu : 2^m \to R_+ \) be a probability distribution on the subsets of the set \( [m] = \{1, 2, \ldots, m\} \). In particular, we assume that \( \mu(.) \) is nonnegative and,
\[ \sum_{S \subseteq [m]} \mu(S) = 1. \]

We assign a multi-affine polynomial with variables \( z_1, \ldots, z_m \) to \( \mu \),
\[ g_\mu(z) = \sum_{S \subseteq [m]} \mu(S) \cdot z^S, \]
where for a set \( S \subseteq [m], z^S = \prod_{i \in S} z_i \). The polynomial \( g_\mu \) is also known as the generating polynomial of \( \mu \). We say \( \mu \) is a homogeneous probability distribution if \( g_\mu \) is a homogeneous polynomial.

We say that \( \mu \) is a strongly Rayleigh distribution if \( g_\mu \) is a real stable polynomial. See Subsection 2.2 for the definition of real stability. Strongly Rayleigh measures are introduced and deeply studied in the seminal work of Borcea, Brändén and Liggett [BBL09]. They are natural generalizations of product distributions and cover several interesting families of probability distributions including determinantal measures and random spanning tree distributions. We refer interested readers to [OSS11, PP14] for applications of these probability measures.

Our main theorem extends Theorem 1.1 to families of vectors assigned to the elements of a strongly Rayleigh distribution. This can be seen as a generalization because independent distributions are special classes of strongly Rayleigh measures. To state the main theorem we need another definition. The marginal probability of an element \( i \) with respect to a probability distribution, \( \mu \), is the probability that \( i \) is in a sample of \( \mu \),
\[ P_{S \sim \mu} \{ i \in S \} = \frac{\partial_{z_i} g_\mu(z)}{\mid \mid z_1 = \ldots = z_m = 1 \mid \mid}. \]

**Theorem 1.2 (Main).** Let \( \mu \) be a homogeneous strongly Rayleigh probability distributions on \( [m] \) such that the marginal probability of each element is at most \( \epsilon_1 \), and let \( v_1, \ldots, v_m \in \mathbb{R}^d \) be vectors such that
\[ \sum_{i=1}^{m} v_i v_i^T = I, \]
and for all \( i \), \( \|v_i\|^2 \leq \epsilon_2 \). Then,
\[ P_{S \sim \mu} \left[ \left\| \sum_{i \in S} v_i v_i^T \right\| \leq 4(\epsilon_1 + \epsilon_2) + 2(\epsilon_1 + \epsilon_2)^2 \right] > 0. \]

We expect to see several applications of the above theorem that are not realizable by the original proof of [MSS13b]. In the following section we describe our main motivation for studying the above statement, which is to design approximation algorithms for the Asymmetric Traveling Salesman Problem (ATSP).
1.1 Spectrally Thin Trees

For a graph \( G = (V, E) \), the Laplacian of \( G \), \( L_G \), is defined as follows: For a vertex \( i \in V \) let \( 1_i \in \mathbb{R}^V \) be the vector that is one at \( i \) and zero everywhere else. Fix an arbitrary orientation on the edges of \( E \) and let \( b_e = 1_i - 1_j \) for an edge \( e \) oriented from \( i \) to \( j \). Then,

\[
L_G = \sum_{e \in E} b_e b_e^\top.
\]

We use \( L_G^\dagger \) to denote the pseudo-inverse of \( L_G \). Also, for a set \( T \subseteq E \), we write

\[
L_T = \sum_{e \in T} b_e b_e^\top.
\]

We say a spanning tree \( T \) is \( \alpha \)-spectrally thin with respect to \( G \) if

\[
L_T \preceq \alpha \cdot L_G.
\]

In our recent work we show that the existence of spectrally thin trees plays an important role in bounding the integrality gap of the natural linear programming relaxation of the Asymmetric TSP \([AO14]\). Given a graph \( G = (V, E) \), Harvey and Olver \([HO14]\) employ a recursive application of \([MSS13b]\) and show that if for all edges \( e \in E \), \( b_e^\top L_G^\dagger b_e \leq \alpha \), then \( G \) has an \( \alpha \)-spectrally thin tree. Unfortunately, this condition is not enough to analyze the integrality gap of the LP relaxation of ATSP \([AO14]\).

Instead, one needs a weaker condition: Suppose there is a set \( F \subseteq E \) such that \((V, F)\) is \( k \)-edge connected\(^1\), and that for all \( e \in F \), \( b_e^\top L_G^\dagger b_e \leq \alpha \). Can we say that \( G \) has a \( C \cdot \max\{\alpha, 1/k\} \)-spectrally thin tree for a universal constant \( C \)?

In this section we use our main theorem to answer the above question affirmatively. This gives an improved upper bound on the integrality gap of the LP relaxation of ATSP. We refer interested readers to \([AO14]\) for more details. Note that the above question cannot be answered by Theorem 1.1. One can use Theorem 1.1 to show that the set \( F \) can be partitioned into two sets \( F_1, F_2 \) such that each \( F_i \) is \( 1/2 + O(\alpha) \)-spectrally thin, but Theorem 1.1 gives no guarantee on the connectivity of \( F_i \)’s. On the other hand, once we apply our main theorem to a strongly Rayleigh distribution supported on connected subgraphs of \( G \), e.g. the spanning trees of \( G \), we get connectivity for free.

First, we prove the following theorem and then we use it to answer the above question.

**Theorem 1.3.** Given a graph \( G = (V, E) \) and \( F \subseteq E \) such that \((V, F)\) is \( k \)-edge connected, if for \( \epsilon > 0 \), we assign vectors \( v_e \) to each edge \( e \in F \) such that

\[
\sum_{e \in F} v_e v_e^\top \preceq I,
\]

and for all \( e \in F \), \( \|v_e\|^2 \leq \epsilon_1 \), then \( G \) has a spanning tree \( T \) such that

\[
\left\| \sum_{e \in T} v_e v_e^\top \right\| \leq O(\epsilon + 1/k).
\]

\(^1\)Recall that a graph is \( k \)-edge connected if it has at least \( k \) edge in any cut.
Proof. To apply our main theorem first we need to construct a strongly Rayleigh distribution of spanning trees of $G$ such that the marginal probability of each edge is “small”.

For $\gamma : E \to \mathbb{R}$, a $\gamma$-uniform random spanning tree distribution is a distribution on spanning trees of $G$ such that for each tree $T$,

$$
\mathbb{P}_\mu [T] \propto \prod_{e \in T} \exp(\gamma_e).
$$

Borcea, Brändén, and Liggett showed that any $\gamma$-uniform distribution of spanning trees is strongly Rayleigh [BBL09]. So, it is enough to find a $\gamma$-uniform random spanning tree distribution such that the marginal probability of each edge is small. Asadpour, Goemans, Madry, the second author, and Saberi [AGM+10] proved a necessary and sufficient condition for the existence of such a distribution. Recall that the spanning tree polytope of $G$ is the convex hull of the characteristic vectors of spanning trees of $G$.

**Theorem 1.4** ([AGM+10, SV14]). For any graph $G = (V, E)$ and any point $x$ in the interior of the spanning tree polytope of $G$, there is a $\gamma$-uniform random spanning tree distribution $\mu$, such that for any edge $e \in E$, $\mathbb{P}_{T \sim \mu} [e \in T] = z_e$.

Nash-Williams [NW61] proved that any $k$-edge connected graph has $k/2$ disjoint spanning trees. Let $T_1, \ldots, T_{k/2}$ be $k/2$ disjoint spanning trees of the given $k$-edge connected graph, $G$. Define $x : E \to \mathbb{R}_+$ as follows: $x_e = 2/k$ for $e \in \bigcup_{i=1}^{k/2} T_i$ and $x_e = 0$ otherwise. By definition, $x$ is in the spanning tree polytope of $G$, so a slight perturbation of $x$ is in the interior of the spanning tree polytope. By the above theorem there is a $\gamma$-uniform distribution $\mu$ such that for each edge $e \in E$,

$$
\mathbb{P}_\mu [e] \leq \frac{2}{k-1}.
$$

To apply **Theorem 1.2** we just need to add a set of vectors $\{v_e\}_{e \in E'}$ of squared norm at most $\epsilon$ such that

$$
\sum_{e \in E'} v_e v_e^T = I - \sum_{e \in E} v_e v_e^T.
$$

This is always possible because the RHS of the above identity is positive semidefinite. Note that any $e \in E'$ has zero marginal probability in $\mu$. Applying **Theorem 1.2** to $\mu$ and vectors $\{v_e\}_{e \in E \cup E'}$, there is a spanning tree $T$ such that

$$
\sum_{e \in T} v_e v_e^T \leq 4\alpha + 2\alpha^2,
$$

for $\alpha = \epsilon + 2/(k-1)$ as desired. \qed

The following corollary is immediate.

**Corollary 1.5.** Given a graph $G = (V, E)$ and a set $F \subseteq E$ such that $(V, F)$ is $k$-edge connected, if for $\epsilon > 0$ and any edge $e \in F$, $b_e L_G^\dagger b_e < \epsilon$, then $G$ has a $O(1/k + \epsilon)$ spectrally thin tree.

**Proof.** Let $L_G^{1/2}$ be the square root of $L_G$. For all $e \in F$, let

$$
v_e = L_G^{1/2} b_e.
$$
Then,
\[
\sum_{e \in F} v_e v_e^\top = L_G^{1/2} \left( \sum_{e \in F} v_e v_e^\top \right) L_G^{1/2} = L_G^{1/2} L_F L_G^{1/2} \leq I.
\]
Therefore, by Theorem 1.3 the graph \((V, F)\) has a spanning tree \(T\) such that
\[
\left\| \sum_{e \in T} v_e v_e^\top \right\| \leq \alpha,
\]
for \(\alpha = O(\epsilon + 1/k)\). Fix an arbitrary vector \(y \in \mathbb{R}^V\). We show that
\[
y^\top L_T y \leq \alpha \cdot y^\top L_G y.
\]
By (2) for any \(x \in \mathbb{R}^V\),
\[
x^\top \left( \sum_{e \in T} v_e v_e^\top \right) x \leq \alpha \cdot \|x\|^2.
\]
Let \(x = L_G^{1/2} y\), we get
\[
y^\top L_G^{1/2} \left( L_G^{1/2} \sum_{e \in T} b_e b_e^\top L_G^{1/2} \right) L_G^{1/2} y \leq \alpha \cdot y^\top L_G y.
\]
The above is the same as (3) and we are done.

1.2 Proof Overview

We build on the method of interlacing polynomials of [MSS13a, MSS13b]. Recall that an interlacing family of polynomials has the property that there is always a polynomial whose largest root is at most the largest root of the sum of the polynomials in the family. In Section 3 we show that for any set of vectors assigned to the elements of a homogeneous strongly Rayleigh measure, the characteristic polynomials of natural quadratic forms associated with the samples of the distribution form an interlacing family.

Analogously to [MSS13b], we define a *mixed characteristic polynomial* as the weighted average of the characteristic polynomials of the natural quadratic forms associated to the samples of the strongly Rayleigh distribution, where the weight of each polynomial is proportional to the probability of the corresponding sample set in the distribution. We show that there is a sample of the distribution such that the largest root of its characteristic polynomial is at most the largest root of the mixed characteristic polynomial.

In Section 4 we extend the multivariate barrier argument of [MSS13b] to upper-bound the largest root of our mixed characteristic polynomial. Unlike [MSS13b], here we need to prove an upper bound on the largest root of the mixed characteristic polynomial which is very close to zero. It turns out that the original idea of [BSS14] that studies the behavior of the roots of a polynomial \(p(x)\) under the operator \(1 - \partial / \partial x\) cannot establish upper bounds that are less than one. Therefore, we study the behavior of the roots of \(p(z)\) under the operators \(1 - \partial / \partial z^2\). Similar to [BSS14, MSS13b]
we keep track of an upper bound on the roots of a polynomial, along with a measure of how far above the roots this upper bound is. The $1 - \partial / \partial z_i^2$ operators allow us to impose very small shifts on this upper bound assuming it is sufficiently far from the roots. This, together with the observation that at the start of the argument the upper bound is close to zero, completes our proof.

2 Preliminaries

We adopt a notation similar to [MSS13b]. We write $\binom{[m]}{k}$ to denote the collection of subsets of $[m]$ with exactly $k$ elements. We write $2^{[m]}$ to denote the family of all subsets of the set $[m]$. We write $\partial z_i$ to denote the operator that performs partial differentiation with respect to $z_i$. We use $\|v\|$ to denote the Euclidean 2-norm of a vector $x$. For a matrix $M$, we write $\|M\| = \max_{\|x\|=1} \|Mx\|$ to denote the operator norm of $M$. We use $\mathbf{1}$ to denote the all 1 vector.

2.1 Interlacing Families

We recall the definition of interlacing families of polynomials from [MSS13a], and its main consequence.

Definition 2.1. We say that a real rooted polynomial $g(x) = \alpha_0 \prod_{i=1}^{m-1} (x - \alpha_i)$ interlaces a real rooted polynomial $f(x) = \beta_0 \prod_{i=1}^{m} (x - \beta_i)$ if

$$\beta_1 \leq \alpha_1 \leq \beta_2 \leq \alpha_2 \leq \ldots \leq \alpha_{m-1} \leq \beta_m.$$  

We say that polynomials $f_1, \ldots, f_k$ have a common interlacing if there is a polynomial $g$ such that $g$ interlaces all $f_i$. The following lemma is proved in [MSS13a].

Lemma 2.2. Let $f_1, \ldots, f_k$ be polynomials of the same degree that are real rooted and have positive leading coefficients. Define

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

If $f_1, \ldots, f_k$ have a common interlacing, then there is an $i$ such that the largest root of $f_i$ is at most the largest root of $f_\emptyset$.

Definition 2.3. Let $\mathcal{F} \subseteq 2^{[m]}$ be nonempty. For any $S \in \mathcal{F}$, let $f_S(x)$ be a real rooted polynomial of degree $d$ with a positive leading coefficient. For $s_1, \ldots, s_k \in \{0, 1\}$ with $k < m$, let

$$\mathcal{F}_{s_1, \ldots, s_k} := \{ S \in \mathcal{F} : i \in S \Leftrightarrow s_i = 1 \}.$$  

Note that $\mathcal{F} = \mathcal{F}_\emptyset$. Define

$$f_{s_1, \ldots, s_k} = \sum_{S \in \mathcal{F}_{s_1, \ldots, s_k}} f_S,$$  

and

$$f_\emptyset = \sum_{S \in \mathcal{F}} f_S.$$  

We say polynomials $\{f_S\}_{S \in \mathcal{F}}$ form an interlacing family if for all $0 \leq k < m$ and all $s_1, \ldots, s_k \in \{0, 1\}$ the following holds: If both of $\mathcal{F}_{s_1, \ldots, s_k, 0}$ and $\mathcal{F}_{s_1, \ldots, s_k, 1}$ are nonempty, $f_{s_1, \ldots, s_k, 0}$ and $f_{s_1, \ldots, s_k, 1}$ have a common interlacing.
The following is analogous to [MSS13b, Thm 3.4].

**Theorem 2.4.** Let $\mathcal{F} \subseteq 2^m$ and let $\{f_S\}_{S \in \mathcal{F}}$ be an interlacing family of polynomials. Then, there exists $S \in \mathcal{F}$ such that the largest root of $f(S)$ is at most the largest root of $f_\emptyset$.

**Proof.** We prove by induction. Assume that for some choice of $s_1, \ldots, s_k \in \{0, 1\}$ (possibly with $k = 0$), $\mathcal{F}_{s_1, \ldots, s_k}$ is nonempty and the largest root of $f_{s_1, \ldots, s_k}$ is at most the largest root of $f_\emptyset$. If $\mathcal{F}_{s_1, \ldots, s_k, 0} = \emptyset$, then $f_{s_1, \ldots, s_k} = f_{s_1, \ldots, s_k, 1}$, so we let $s_{k+1} = 1$ and we are done. Similarly, if $\mathcal{F}_{s_1, \ldots, s_k, 1} = \emptyset$, then we let $s_{k+1} = 0$ and we are done with the induction. If both of these sets are nonempty, then $f_{s_1, \ldots, s_k, 0}$ and $f_{s_1, \ldots, s_k, 1}$ have a common interlacing. So, by Lemma 2.2, for some choice of $s_{k+1} \in \{0, 1\}$, the largest root of $f_{s_1, \ldots, s_{k+1}}$ is at most the largest root of $f_\emptyset$. □

We use the following lemma which appeared as Theorem 2.1 of [Ded92] to prove that a certain family of polynomials that we construct in Section 3 form an interlacing family.

**Lemma 2.5.** 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 convex combinations $\lambda_i \geq 0$, $\sum_{i=1}^k \lambda_i = 1$.

### 2.2 Stable Polynomials

Stable polynomials are natural multivariate generalizations of real-rooted univariate polynomials. For a complex number $z$, let $\text{Im}(z)$ denote the imaginary part of $z$. We say a polynomial $p(z_1, \ldots, z_m) \in \mathbb{C}[z_1, \ldots, z_m]$ is stable if whenever $\text{Im}(z_i) > 0$ for all $1 \leq i \leq m$, $p(z_1, \ldots, z_m) \neq 0$. We say $p(\cdot)$ is real stable, if it is stable and all of its coefficients are real. It is easy to see that any univariate polynomial is real stable if and only if it is real rooted.

One of the most interesting classes of real stable polynomials is the class of determinant polynomials as observed by Borcea and Brändén [BB08].

**Theorem 2.6.** For any set of positive semidefinite matrices $A_1, \ldots, A_m$, the following polynomial is real stable:

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

Perhaps the most important property of stable polynomials is that they are closed under several elementary operations like multiplication, differentiation, and substitution. We will use these operations to generate new stable polynomials from the determinant polynomial. The following is proved in [MSS13b].

**Lemma 2.7.** If $p \in \mathbb{R}[z_1, \ldots, z_m]$ is real stable, then so are the polynomials $(1 - \partial_{z_1})p(z_1, \ldots, z_m)$ and $(1 + \partial_{z_1})p(z_1, \ldots, z_m)$.

The following corollary simply follows from the above lemma.

**Corollary 2.8.** If $p \in \mathbb{R}[z_1, \ldots, z_m]$ is real stable, then so is

$$(1 - \partial_{z_1}^2)p(z_1, \ldots, z_m).$$
Proof. First, observe that

\[(1 - \partial_{z_1}^2)p(z_1, \ldots, z_m) = (1 - \partial_{z_1})(1 + \partial_{z_1})p(z_1, \ldots, z_m).\]

So, the conclusion follows from two applications of Lemma 2.7.

The following closure properties are elementary.

**Lemma 2.9.** If \( p \in \mathbb{R}[z_1, \ldots, z_m] \) is real stable, then so is \( p(\lambda \cdot z_1, \ldots, \lambda_m \cdot z_m) \) for real-valued \( \lambda_1, \ldots, \lambda_m > 0 \).

*Proof.* Say \((z_1, \ldots, z_m) \in \mathbb{C}^m\) is a root of \( p(\lambda \cdot z_1, \ldots, \lambda_m \cdot z_m) \). Then \((\lambda \cdot z_1, \ldots, \lambda_m \cdot z_m) \) is a root of \( p(z_1, \ldots, z_m) \). Since \( p \) is real stable, there is an \( i \) such that \( \text{Im}(\lambda_i \cdot z_i) \leq 0 \). But, since \( \lambda_i > 0 \), we get \( \text{Im}(z_i) \leq 0 \), as desired.

**Lemma 2.10.** If \( p \in \mathbb{R}[z_1, \ldots, z_m] \) is real stable, then so is \( p(z_1 + x, \ldots, z_m + x) \) for a new variable \( x \).

*Proof.* Say \((z_1, \ldots, z_m, x) \in \mathbb{C}^m\) is a root of \( p(z_1 + x, \ldots, z_m + x) \). Then \((z_1 + x, \ldots, z_m + x) \) is a root of \( p(z_1, \ldots, z_m) \). Since \( p \) is real stable, there is an \( i \) such that \( \text{Im}(z_i + x) \leq 0 \). But, then either \( \text{Im}(x) \leq 0 \) or \( \text{Im}(z_i) \leq 0 \), as desired.

### 2.3 Facts from Linear Algebra

For a Hermitian matrix \( M \in \mathbb{C}^{d \times d} \), we write the characteristic polynomial of \( M \) in terms of a variable \( x \) as

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

We also write the characteristic polynomial in terms of the square of \( x \) as

\[\chi[M](x^2) = \det(x^2I - M).\]

For \( 1 \leq k \leq n \), we write \( \sigma_k(M) \) to denote the sum of all principal \( k \times k \) minors of \( M \), in particular,

\[\chi[M](x) = \sum_{k=0}^{d} x^{d-k}(-1)^k \sigma_k(M).\]

The following lemma follows from the Cauchy-Binet identity. See [MSS13b] for the proof.

**Lemma 2.11.** For vectors \( v_1, \ldots, v_m \in \mathbb{R}^d \) and scalars \( z_1, \ldots, z_m \),

\[
\det \left( xI + \sum_{i=1}^{m} z_i v_i v_i^\top \right) = \sum_{k=0}^{d} x^{d-k} \sum_{S \subseteq \binom{[m]}{k}} z^S \sigma_k \left( \sum_{i \in S} v_i v_i^\top \right).
\]

In particular, for \( z_1 = \ldots = z_m = -1 \),

\[
\det \left( xI - \sum_{i=1}^{m} v_i v_i^\top \right) = \sum_{k=0}^{d} x^{d-k}(-1)^k \sum_{S \subseteq \binom{[m]}{k}} \sigma_k \left( \sum_{i \in S} v_i v_i^\top \right).
\]
The following is Jacboi’s formula for the derivative of the determinant of a matrix.

**Theorem 2.12.** For an invertible matrix $A$ which is a differentiable function of $t$, \( \partial_t \det(A) = \det(A) \cdot \text{Tr}(A^{-1} \partial_t A) \).

**Lemma 2.13.** For an invertible matrix $A$ which is a differentiable function of $t$, \( \frac{\partial A^{-1}}{\partial t} = -A^{-1}(\partial_t A)A^{-1} \).

**Proof.** Differentiating both sides of the identity $A^{-1}A = I$ with respect to $t$, we get \( A^{-1} \frac{\partial A}{\partial t} + \frac{\partial A^{-1}}{\partial t} A = 0 \). Rearranging the terms and multiplying with $A^{-1}$ gives the lemma’s conclusion.

The following two standard facts about trace will be used throughout the paper. First, for $A \in \mathbb{R}^{k \times n}$ and $B \in \mathbb{R}^{n \times k}$, \( \text{Tr}(AB) = \text{Tr}(BA) \).

Secondly, for positive semidefinite matrices $A, B$ of the same dimension, \( \text{Tr}(AB) \geq 0 \).

Also, we use the fact that for any positive semidefinite matrix $A$ and any Hermitian matrix $B$, $BAB$ is positive semidefinite.

### 3 The Mixed Characteristic Polynomial

For a probability distribution $\mu$, let $d_\mu$ be the degree of the polynomial $g_\mu$.

**Theorem 3.1.** For $v_1, \ldots, v_m \in \mathbb{R}^d$ and a homogeneous probability distribution $\mu : [m] \rightarrow \mathbb{R}_+$, \[
\mathbb{E}_{S \sim \mu} x^{d_\mu - d} \left( \sum_{i \in S} 2v_i v_i^\top \right) (x^2) = \prod_{i=1}^m (1 - \partial_{z_i}^2) \left( g_\mu(x1 + z) \cdot \det \left( xI + \sum_{i=1}^m z_i v_i v_i^\top \right) \right) \bigg|_{z_1 = \ldots = z_m = 0}.
\]

We call the polynomial $\mathbb{E}_{S \sim \mu} x^{\sum_{i \in S} 2v_i v_i^\top}(x^2)$ the mixed characteristic polynomial and we denote it by $\mu[v_1, \ldots, v_m](x)$.

**Proof.** For $S \subseteq [m]$, let $z^{2S} = \prod_{i \in S} z_i^2$. By Lemma 2.11, the coefficient of $z^{2S}$ in \[g_\mu(x1 + z) \cdot \det(xI + \sum_{i=1}^m z_i v_i v_i^\top)\]
is equal to \[
\left( \prod_{i \in S} \partial_{z_i}^2 \right) \left( g_\mu(x1 + z) \cdot \det \left( xI + \sum_{i=1}^m z_i v_i v_i^\top \right) \right) \bigg|_{z_1 = \ldots = z_m = 0}.
\]
Each of the two polynomials $g_\mu(x_1 + z)$ and $\det(xI + \sum_{i=1}^m z_i v_i v_i^\top)$ is multi-linear in $z_1, \ldots, z_m$. Therefore, for $k = |S|$, the above is equal to

$$2^k \cdot \left( \prod_{i \in S} \partial_{z_i} \right) g_\mu(x_1 + z) \bigg|_{z_1 = \ldots = z_m = 0} \cdot \left( \prod_{i \in S} \partial_{z_i} \right) \det \left( xI + \sum_{i=1}^m z_i v_i v_i^\top \right) \bigg|_{z_1 = \ldots = z_m = 0}. \quad (5)$$

Since $g_\mu$ is a homogeneous polynomial of degree $d_\mu$, the first term in the above is equal to

$$x^{d_\mu - k} \mathbb{P}_{T \sim \mu} [S \subseteq T].$$

And, by Lemma 2.11, the second term of (5) is equal to

$$x^{d - k} \sigma_k \left( \sum_{i \in S} z_i v_i v_i^\top \right).$$

Applying the above identities for all $S \subseteq [m],$

$$\prod_{i=1}^m (1 - \partial_{z_i}^2) \left( g_\mu(x_1 + z) \cdot \det \left( xI + \sum_{i=1}^m z_i v_i v_i^\top \right) \right) \bigg|_{z_1 = \ldots = z_m = 0}$$

$$= \sum_{k=0}^d (-1)^k \sum_{S \subseteq [m]} \left( \prod_{i \in S} \partial_{z_i}^2 \right) \left( g_\mu(x_1 + z) \cdot \det \left( xI + \sum_{i=1}^m z_i v_i v_i^\top \right) \right) \bigg|_{z_1 = \ldots = z_m = 0}$$

$$= \sum_{k=0}^d (-1)^k x^{d_\mu + d - 2k} \mathbb{P}_{S \sim \mu} \mathbb{P}_{T \sim \mu} [S \subseteq T] \cdot \sigma_k \left( \sum_{i \in S} z_i v_i v_i^\top \right)$$

$$= x^{d_\mu - d} \mathbb{E}_{S \sim \mu} \chi \left[ \sum_{i \in S} 2v_i v_i^\top \right] (x^2).$$

The last identity uses Lemma 2.11.

**Corollary 3.2.** If $\mu$ is a strongly Rayleigh probability distribution, then the mixed characteristic polynomial is real-rooted.

**Proof.** First, by Theorem 2.6,

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

is real stable. Since $\mu$ is strongly Rayleigh, $g_\mu(z)$ is real stable. So, by Lemma 2.10, $g_\mu(x_1 + z)$ is real stable. The product of two real stable polynomials is also real stable, so

$$g_\mu(x_1 + z) \cdot \det \left( xI + \sum_{i=1}^m z_i v_i v_i^\top \right)$$

is real stable. Corollary 2.8 implies that

$$\prod_{i=1}^m (1 - \partial_{z_i}^2) \left( g_\mu(x_1 + z) \cdot \det \left( xI + \sum_{i=1}^m z_i v_i v_i^\top \right) \right)$$

is real stable. \qed
is real stable as well. Wagner [Wag11, Lemma 2.4(d)] tells us that real stability is preserved under setting variables to real numbers, so

\[
\prod_{i=1}^{m} (1 - \partial^2_{z_i}) \left( g_\mu(x1 + z) \cdot \det \left( xI + \sum_{i=1}^{m} z_i v_i v_i^\top \right) \right) \bigg|_{z_1=\ldots=z_m=0}
\]

is a univariate real-rooted polynomial. The mixed characteristic polynomial is equal to the above polynomial up to a term \(x^{d_\mu-d}\). So, the mixed characteristic polynomial is also real rooted. \(\square\)

Now, we use the real-rootedness of the mixed characteristic polynomial to show that the characteristic polynomials of the set of vectors assigned to any set \(S\) with nonzero probability in \(\mu\) form an interlacing family. For a homogeneous strongly Rayleigh measure \(\mu\), let

\[
\mathcal{F} = \{S : \mu(S) > 0\},
\]

and for \(s_1, \ldots, s_k \in \{0,1\}\) let \(\mathcal{F}_{s_1, \ldots, s_k}\) be as defined in Definition 2.3. For any \(S \in \mathcal{F}\), let

\[
q_S(x) = \mu(S) \cdot \chi \left[ \sum_{i \in S} 2v_i v_i^\top \right] (x^2).
\]

**Theorem 3.3.** The polynomials \(\{q_S\}_{S \in \mathcal{F}}\) form an interlacing family.

**Proof.** For \(1 \leq k \leq m\) and \(s_1, \ldots, s_k \in \{0,1\}\), let \(\mu_{s_1, \ldots, s_k}\) be \(\mu\) conditioned on the sets \(S \in \mathcal{F}_{s_1, \ldots, s_k}\), i.e., \(\mu\) conditioned on \(i \in S\) for all \(i \leq k\) where \(s_i = 1\) and \(i \notin S\) for all \(i \leq k\) where \(s_i = 0\). We inductively write the generating polynomial of \(\mu_{s_1, \ldots, s_k}\) in terms of \(g_\mu\). Say we have written \(g_{\mu_{s_1, \ldots, s_k}}\) in terms of \(g_\mu\). Then, we can write,

\[
g_{\mu_{s_1, \ldots, s_k,1}}(z) = \frac{z_{k+1} \cdot \partial_{z_{k+1}} g_{\mu_{s_1, \ldots, s_k}}(z)}{\partial_{z_{k+1}} g_{\mu_{s_1, \ldots, s_k}}(z)} \bigg|_{z_i=1} = \frac{g_{\mu_{s_1, \ldots, s_k}}(z)}{g_{\mu_{s_1, \ldots, s_k}}(z) \bigg|_{z_{k+1}=0, z_i=1 \text{ for } i \neq k+1}}.
\]

Note that the denominators of both equations are just normalizing constants. The above polynomials are well defined if the normalizing constants are nonzero, i.e., if the set \(\mathcal{F}_{s_1, \ldots, s_k, s_{k+1}}\) is nonempty. Since the real stable polynomials are closed under differentiation and substitution, for any \(1 \leq k \leq m\), and \(s_1, \ldots, s_k \in \{0,1\}\), if \(g_{\mu_{s_1, \ldots, s_k}}\) is well defined, it is real stable, so \(\mu_{s_1, \ldots, s_k}\) is a strongly Rayleigh distribution.

Now, for \(s_1, \ldots, s_k \in \{0,1\}\), let

\[
q_{s_1, \ldots, s_k}(x) = \sum_{S \in \mathcal{F}_{s_1, \ldots, s_k}} q_S(x).
\]

Since \(\mu_{s_1, \ldots, s_k}\) is strongly Rayleigh, by Corollary 3.2, \(q_{s_1, \ldots, s_k}(x)\) is real rooted.

By Lemma 2.5, to prove the theorem it is enough to show that if \(\mathcal{F}_{s_1, \ldots, s_k,0}\) and \(\mathcal{F}_{s_1, \ldots, s_k,1}\) are nonempty, then for any \(0 < \lambda < 1\),

\[
\lambda \cdot q_{s_1, \ldots, s_k,1}(x) + (1 - \lambda) \cdot q_{s_1, \ldots, s_k,0}(x)
\]
is real rooted. Equivalently, by Corollary 3.2, it is enough to show that for any $0 < \lambda < 1$,
\[
\lambda \cdot g_{\mu_1, \ldots, \mu_k}(z) + (1 - \lambda) \cdot g_{\mu_1, \ldots, \mu_k, 0}(z)
\]
is real stable. Let us write,
\[
g_{\mu_1, \ldots, \mu_k}(z) = z_{k+1} \cdot \frac{\partial}{\partial z_{k+1}} g_{\mu_1, \ldots, \mu_k}(z) + g_{\mu_1, \ldots, \mu_k}(z) \bigg|_{z_{k+1} = 0}
\]
for some $\alpha, \beta > 0$. The second identity follows by (6) and (7). Let $\lambda_{k+1} > 0$ such that
\[
\frac{\lambda_{k+1} \cdot \alpha}{\lambda} = \frac{\beta}{1 - \lambda}.
\]
Since $g_{\mu_1, \ldots, \mu_k}$ is real stable, by Lemma 2.9
\[
g_{\mu_1, \ldots, \mu_k}(z_1, \ldots, \lambda_{k+1} \cdot z_{k+1}, z_{k+2}, \ldots, z_m)
\]
is real stable. But, by (9) the above polynomial is just a multiple of (8). So, (8) is real stable. $\square$

4 An Extension of [MSS13b] Multivariate Barrier Argument

In this section we upper-bound the roots of the mixed characteristic polynomial in terms of the marginal probabilities of elements of $[m]$ in $\mu$ and the maximum of the squared norm of vectors $v_1, \ldots, v_m$.

**Theorem 4.1.** Given vectors $v_1, \ldots, v_m \in \mathbb{R}^d$, and a homogeneous probability distribution $\mu : [m] \rightarrow \mathbb{R}_+$, such that the marginal probability of each element $i \in [m]$ is at most $\epsilon_1$, $\sum_{i=1}^m v_i v_i^\top = I$ and $\|v_i\|^2 \leq \epsilon_2$, the largest root of $\mu[v_1, \ldots, v_m](x)$ is at most $4(2\epsilon + \epsilon^2)$, where $\epsilon = \epsilon_1 + \epsilon_2$.

First, similar to [MSS13b] we derive a slightly different expression.

**Lemma 4.2.** For any probability distribution $\mu$ and vectors $v_1, \ldots, v_m \in \mathbb{R}^d$ such that $\sum_{i=1}^m v_i v_i^\top = I$,
\[
x^{d_{\mu} - d}[v_1, \ldots, v_m](x) = \prod_{i=1}^m (1 - \partial_{y_i}^2) \left( g_{\mu}(y) \cdot \det \left( \sum_{i=1}^m y_i v_i v_i^\top \right) \right) \bigg|_{y_1 = \ldots = y_m = x}.
\]

**Proof.** This is because for any differentiable function $f$, $\partial_{y_i} f(y_i)|_{y_i = z_i + x} = \partial_{z_i} f(z_i + x)$. $\square$

Let
\[
Q(y_1, \ldots, y_m) = \prod_{i=1}^m (1 - \partial_{y_i}^2) \left( g_{\mu}(y) \cdot \det \left( \sum_{i=1}^m y_i v_i v_i^\top \right) \right).
\]

Then, by the above lemma, the maximum root of $Q(x, \ldots, x)$ is the same as the maximum root of $\mu[v_1, \ldots, v_m](x)$. In the rest of this section we upper-bound the maximum root of $Q(x, \ldots, x)$.

It directly follows from the proof of Theorem 5.1 in [MSS13b] that the maximum root of $Q(x, \ldots, x)$ is at most $(1 + \sqrt{\epsilon})^2$. But, in our setting, any upper-bound that is more than 1 obviously holds, as for any $S \subseteq [m]$,
\[
\left| \sum_{i=1}^m v_i v_i^\top \right| \leq 1.
\]
The main difficulty that we are facing is to prove an upper-bound of $O(\epsilon)$ on the maximum root of $Q(x, \ldots, x)$.

We use an extension of the multivariate barrier argument of [MSS13b] to upper-bound the maximum root of $Q$. We manage to prove a significantly smaller upper-bound because we apply $1 - \partial^2_{x_i}$ operators as opposed to the $1 - \partial_{x_i}$ operators used in [MSS13b]. This allows us to impose significantly smaller shifts on the barrier upper-bound in our inductive argument.

**Definition 4.3.** For a multivariate polynomial $p(z_1, \ldots, z_m)$, we say $z \in \mathbb{R}^m$ is above all roots of $p$ if for all $t \in \mathbb{R}^m_+$,

$$p(z + t) > 0.$$  

We use $\text{Ab}_p$ to denote the set of points which are above all roots of $p$.

We use the same barrier function defined in [MSS13b].

**Definition 4.4.** For a real stable polynomial $p$, and $z \in \text{Ab}_p$, the barrier function of $p$ in direction $i$ at $z$ is

$$\Phi^i_p(z) := \frac{\partial z_i p(z)}{p(z)} = \partial z_i \log p(z).$$

To analyze the rate of change of the barrier function with respect to the $1 - \partial^2_{z_i}$ operator, we need to work with the second derivative of $p$ as well. We define,

$$\Psi^i_p(z) := \frac{\partial^2 z_i p(z)}{p(z)}.$$

Equivalently, for a univariate restriction $q_{z,i}(t) = p(z_1, \ldots, z_{i-1}, t, z_{i+1}, \ldots, z_m)$, with real roots $\lambda_1, \ldots, \lambda_r$ we can write,

$$\Phi^i_p(z) = \frac{q'_{z,i}(z_i)}{q_{z,i}(z_i)} = \sum_{j=1}^r \frac{1}{z_i - \lambda_j},$$

$$\Psi^i_p(z) = \frac{q''_{z,i}(z_i)}{q_{z,i}(z_i)} = \sum_{1 \leq j < k \leq r} \frac{2}{(z_i - \lambda_j)(z_i - \lambda_k)}.$$

The following lemma is immediate from the above definition.

**Lemma 4.5.** If $p$ is real stable and $z \in \text{Ab}_p$, then for all $i \leq m$,

$$\Psi^i_p(z) \leq \Phi^i_p(z)^2.$$

**Proof.** Since $z \in \text{Ab}_p$, $z_i > \lambda_j$ for all $1 \leq j \leq r$, so,

$$\Phi^i_p(z)^2 - \Psi^i_p(z) = \left( \sum_{j=1}^r \frac{1}{z_i - \lambda_j} \right)^2 - \sum_{1 \leq j < k \leq r} \frac{2}{(z_i - \lambda_j)(z_i - \lambda_k)} = \sum_{j=1}^r \frac{1}{(z_i - \lambda_j)^2} > 0.$$

The following monotonicity and convexity properties of the barrier functions are proved in [MSS13b].
Lemma 4.6. Suppose \( p(.) \) is a real stable polynomial and \( z \in \mathbb{A} p \). Then, for all \( i, j \leq m \) and \( \delta \geq 0 \),

\[
\Phi^i_p(z + \delta 1_j) \leq \Phi^i_p(z) \quad \text{and}, \quad (\text{monotonicity}) \tag{10}
\]

\[
\Phi^i_p(z + \delta 1_j) \leq \Phi^i_p(z) + \delta \cdot \partial_{z_j} \Phi^i_p(z + \delta 1_j) \quad (\text{convexity}). \tag{11}
\]

Recall that the purpose of the barrier functions \( \Phi^i_p \) is to allow us to reason about the relationship between \( \mathbb{A} p \) and \( \mathbb{A} p - \partial^2_{z_j} p \); the monotonicity property and Lemma 4.5 imply the following lemma.

Lemma 4.7. If \( p \) is real stable and \( z \in \mathbb{A} p \) is such that \( \Phi^i_p(z) < 1 \), then \( z \in \mathbb{A} p - \partial^2_{z_j} p \).

Proof. Fix a nonnegative vector \( t \). Since \( \Phi \) is nonincreasing in each coordinate,

\[
\Phi^i_p(z + t) \leq \Phi^i_p(z) < 1.
\]

Since \( z + t \in \mathbb{A} p \), by Lemma 4.5,

\[
\Psi^i_p(z + t) \leq \Phi^i_p(z + t)^2 < 1.
\]

Therefore,

\[
\partial^2_{z_i} p(z + t) < p(z + t) \Rightarrow (1 - \partial^2_{z_i} p)(z + t) > 0,
\]

as desired. \( \square \)

We use an inductive argument similar to [MSS13b]. We argue that when we apply each operator \( (1 - \partial^2_{z_j}) \), the barrier functions, \( \Phi^i_p(z) \), do not increase by shifting the upper bound along the direction \( 1_j \). As we would like to prove a significantly smaller upper bound on the maximum root of the mixed characteristic polynomial, we may only shift along direction \( 1_j \) by a small amount. In the following lemma we show that when we apply the \( (1 - \partial^2_{z_j}) \) operator we only need to shift the upper bound proportional to \( \Phi^i_p(z) \) along the direction \( 1_j \).

Lemma 4.8. Suppose that \( p(z_1, \ldots , z_m) \) is real stable and \( z \in \mathbb{A} p \). If for \( \delta > 0 \),

\[
\frac{2}{\delta} \Phi^i_p(z) + \Phi^i_p(z) \leq 1,
\]

then, for all \( i \),

\[
\Phi^i_{p - \partial^2_{z_j} p}(z + \delta \cdot 1_j) \leq \Phi^i_p(z).
\]

To prove the above lemma we first need to prove a technical lemma to upper-bound \( \frac{\partial_{z_i} \Psi^i_p(z)}{\partial_{z_j} \Phi^i_p(z)} \). We use the following characterization of the bivariate real stable polynomials proved by Lewis, Parrilo, and Ramana [LPR05]. The following form is stated in [BB10, Cor 6.7].

Lemma 4.9. If \( p(z_1, z_2) \) is a bivariate real stable polynomial of degree \( d \), then there exist \( d \times d \) positive semidefinite matrices \( A, B \) and a Hermitian matrix \( C \) such that

\[
p(z_1, z_2) = \pm \det(z_1 A + z_2 B + C).
\]
**Lemma 4.10.** Suppose that \( p \) is real stable and \( z \in \mathbf{Ab}_p \), then for all \( i, j \leq m \),

\[
\frac{\partial_z \Psi_p^j(z)}{\partial_z \Phi_p^j(z)} \leq 2\Phi_p^j(z).
\]

**Proof.** If \( i = j \), then we consider the univariate restriction \( q_{z,i}(z_i) = \prod_{k=1}^r (z_i - \lambda_k) \). Then,

\[
\frac{\partial_z \sum_{1 \leq k \leq r} \frac{2}{(z_i - \lambda_k)^2}}{\partial_z \sum_{k=1}^r (z_i - \lambda_k)} = \frac{\sum_{k \neq \ell} (z_i - \lambda_k)^2}{\sum_{k=1}^r (z_i - \lambda_k)^2} \leq \sum_{\ell=1}^r \frac{2}{(z_i - \lambda_\ell)} = 2\Phi_p^j(z).
\]

The inequality uses the assumption that \( z \in \mathbf{Ab}_p \).

If \( i \neq j \), we fix all variables other than \( z_i, z_j \) and we consider the bivariate restriction

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

By **Lemma 4.9**, there are Hermitian positive semidefinite matrices \( B_i, B_j \), and a Hermitian matrix \( C \) such that

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

Let \( M = z_i B_i + z_j B_j + C \). Marcus, Spielman, and Srivastava [MSS13b, Lem 5.7] observed that the sign is always positive, that \( B_i + B_j \) is positive definite. In addition, \( M \) is positive definite since \( B_i + B_j \) is positive definite and \( z \in \mathbf{Ab}_p \).

By **Theorem 2.12**, the barrier function in direction \( j \) can be expressed as

\[
\Phi_p^j(z) = \frac{\partial_z \det(M)}{\det(M)} = \frac{\det(M) \text{Tr}(M^{-1}B_j)}{\det(M)} = \text{Tr}(M^{-1}B_j).
\]

By another application of **Theorem 2.12**,\n
\[
\Psi_p^j(z) = \frac{\partial_z^2 \det(M)}{\det(M)} = \frac{\partial_z (\det(M) \text{Tr}(M^{-1}B_j))}{\det(M)} = \frac{\det(M) \text{Tr}(M^{-1}B_j)^2 + \det(M) \text{Tr}((\partial_z M^{-1})B_j)}{\det(M)} = \text{Tr}(M^{-1}B_j)^2 + \text{Tr}(-M^{-1}B_jM^{-1}B_j) = \text{Tr}(M^{-1}B_j)^2 - \text{Tr}((M^{-1}B_j)^2).
\]

The second to last identity uses **Lemma 2.13**. Next, we calculate \( \partial_z \Phi_p^j \) and \( \partial_z \Psi_p^j \). First, by another application of **Lemma 2.13**,\n
\[
\partial_z M^{-1}B_j = -M^{-1}B_iM^{-1}B_j =: L.
\]

Therefore,

\[
\partial_z \Phi_p^j(z) = \partial_z \text{Tr}(M^{-1}B_j) = \text{Tr}(L),
\]

and

\[
\partial_z \Psi_p^j(z) = \partial_z \text{Tr}(M^{-1}B_j)^2 - \partial_z \text{Tr}((M^{-1}B_j)^2) = 2\text{Tr}(M^{-1}B_j) \text{Tr}(L) - \text{Tr}(L(M^{-1}B_j) + (M^{-1}B_j)L) = 2\text{Tr}(M^{-1}B_j) \text{Tr}(L) - 2\text{Tr}(LM^{-1}B_j).
\]
Putting above equations together we get
\[
\frac{\partial_z \psi^j_p(z)}{\partial z_i \phi^j_p(z)} = 2 \frac{\text{Tr}(M^{-1}B_j) \text{Tr}(L) - \text{Tr}(LM^{-1}B_j)}{\text{Tr}(L)} = 2 \frac{\text{Tr}(M^{-1}B_j) - 2 \text{Tr}(LM^{-1}B_j)}{\text{Tr}(L)} = 2\phi^j_p(z) - 2 \frac{\text{Tr}(LM^{-1}B_j)}{\text{Tr}(L)}
\]
where we used (12).

To prove the lemma it is enough to show that \(\frac{\text{Tr}(LM^{-1}B_j)}{\text{Tr}(L)} \geq 0\). We show that both the numerator and the denominator are nonpositive. First,
\[
\text{Tr}(L) = -\text{Tr}(M^{-1}B_i M^{-1}B_j) \leq 0
\]
where we used that \(M^{-1}B_i M^{-1}B_j\) and \(B_i\) are positive semidefinite and the fact that the trace of the product of positive semidefinite matrices is nonnegative. Secondly,
\[
\text{Tr}(LM^{-1}B_j) = \text{Tr}(-M^{-1}B_i M^{-1}B_j M^{-1}B_j) = -\text{Tr}(B_i M^{-1}B_j M^{-1}B_j M^{-1}) \leq 0,
\]
where we again used that \(M^{-1}B_j M^{-1}B_j M^{-1}\) and \(B_i\) are positive semidefinite and the trace of the product of two positive semidefinite matrices is nonnegative.

**Proof of Lemma 4.8.** We write \(\partial_i\) instead of \(\partial_{zi}\) for the ease of notation. First, we write \(\phi^j_p(z) - \partial^2_{zj} p\) in terms of \(\phi^i_p\) and \(\psi^j_p\) and \(\partial_i \psi^j_p\).

\[
\Phi^i_{p-\partial^2_j p} = \frac{\partial_i(p - \partial^2_j p)}{p - \partial^2_j p} = \frac{\partial_i((1 - \psi^j_p)p)}{(1 - \psi^j_p)p} = \frac{(1 - \psi^j_p)(\partial_i p)}{(1 - \psi^j_p)p} + \frac{(\partial_i(1 - \psi^j_p))p}{(1 - \psi^j_p)p} = \phi^j_p - \frac{\partial_i \psi^j_p}{1 - \psi^j_p}. 
\]

We would like to show that \(\Phi^i_{p-\partial^2_j p}(z + \delta 1_j) \leq \Phi^i_p(z)\). Equivalently, it is enough to show that
\[
-\frac{\partial_i \psi^j_p(z + \delta 1_j)}{1 - \psi^j_p(z + \delta 1_j)} \leq \Phi^i_p(z) - \Phi^i_p(z + \delta 1_j) .
\]
By (11) of Lemma 4.6, it is enough to show that
\[
-\frac{\partial_i \psi^j_p(z + \delta 1_j)}{1 - \psi^j_p(z + \delta 1_j)} \leq \delta \cdot (-\partial_j \phi^i_p(z + \delta 1_j)).
\]
By (10) of Lemma 4.6, \( \delta \cdot (-\partial_j \Phi_p(z + \delta 1_j)) > 0 \) so we may divide both sides of the above inequality by this term and obtain

\[
\frac{-\partial_i \Psi_p^j(z + \delta 1_j)}{-\delta \cdot \partial_i \Phi_p(z + \delta 1_j)} \cdot \frac{1}{1 - \Psi_p^j(z + \delta 1_j)} \leq 1,
\]

where we also used \( \partial_j \Phi_p^i = \partial_i \Phi_p^j \). By Lemma 4.10, \( \frac{\partial \Psi_p^j}{\partial \Phi_p^j} \leq 2 \Phi_p^j \). So, we can write,

\[
\frac{2}{\delta} \Phi_p^j(z + \delta 1_j) \cdot \frac{1}{1 - \Psi_p^j(z + \delta 1_j)} \leq 1.
\]

By Lemma 4.5 and (10) of Lemma 4.6,

\[
\Phi_p^j(z + \delta 1_j) \leq \Phi_p^j(z),
\]

\[
\Psi_p^j(z + \delta 1_j) \leq \Phi_p^j(z + \delta 1_j)^2 \leq \Phi_p^j(z)^2.
\]

So, it is enough to show that

\[
\frac{2}{\delta} \Phi_p^j(z) \cdot \frac{1}{1 - \Phi_p^j(z)^2} \leq 1.
\]

Using \( \Phi_p^j(z) < 1 \) we may multiply both sides with \( 1 - \Phi_p^j(z) \) and we obtain,

\[
\frac{2}{\delta} \Phi_p^j(z) + \Phi_p^j(z)^2 \leq 1,
\]

as desired. \( \square \)

Now, we are ready to prove Theorem 4.1.

**Proof of Theorem 4.1.** Let

\[
p(y_1, \ldots, y_m) = g_\mu(y) \cdot \det \left( \sum_{i=1}^m y_i v_i v_i^\top \right).
\]

Set \( \epsilon = \epsilon_1 + \epsilon_2 \) and

\[
\delta = t = \sqrt{2\epsilon + \epsilon^2}.
\]

For any \( z \in \mathbb{R}^m \) with positive coordinates, \( g_\mu(z) > 0 \), and additionally

\[
\det \left( \sum_{i=1}^m z_i v_i v_i^\top \right) > 0.
\]

Therefore, for every \( t > 0 \), \( t1 \in \mathbb{A} \) \( p \).

Now, by Theorem 2.12,

\[
\Phi_p^j(y) = \frac{(\partial_i g_\mu(y)) \cdot \det(\sum_{i=1}^m y_i v_i v_i^\top)}{g_\mu(y) \cdot \det(\sum_{i=1}^m y_i v_i v_i^\top)} + \frac{g_\mu(y) \cdot (\partial_i \det(\sum_{i=1}^m y_i v_i v_i^\top))}{g_\mu(y) \cdot \det(\sum_{i=1}^m y_i v_i v_i^\top)}
\]

\[
= \frac{\partial_i g_\mu(y)}{g_\mu(y)} + \text{Tr} \left( \sum_{i=1}^m y_i v_i v_i^\top \right)^{-1} y_i y_i^\top
\]

17
Therefore, since $g_\mu$ is homogeneous,
\[
\Phi^i_p(t \mathbf{1}) = \frac{1}{t} \cdot \frac{\partial g_\mu(1)}{g_\mu(1)} + \frac{||v_i||^2}{t} = \frac{\mathbb{P}_{S \sim \mu}[i \in S]}{t} + \frac{||v_i||^2}{t} \leq \frac{\epsilon_1}{t} + \frac{\epsilon_2}{t} = \frac{\epsilon}{t}.
\]
The second identity uses (1). Let $\phi = \epsilon/t$. Using $t = \delta$, it follows that
\[
\frac{2}{\delta} \phi + \phi^2 = \frac{2\epsilon}{t^2} + \frac{\epsilon^2}{t^2} = 1.
\]

For $k \in [m]$ define
\[
p_k(y_1, \ldots, y_m) = \prod_{i=1}^k (1 - \partial_{y_i}^2) \left( g_\mu(y) \cdot \det \left( \sum_{i=1}^m y_i v_i v_i^\top \right) \right),
\]
and note that $p_m = Q$. Let $x^0$ be the all-$t$ vector and $x^k$ be the vector that is $t + \delta$ in the first $k$ coordinates and $t$ in the rest. By inductively applying Lemma 4.7 and Lemma 4.8 for any $k \in [m]$, $x^k$ is above all roots of $p_k$ and for all $i$,
\[
\Phi^i_{p_k}(x_k) \leq \phi \Rightarrow \frac{2}{\delta} \Phi^i_{p_k}(x_i) + \Phi^2_{p_k}(x_i) \leq 1.
\]

Therefore, the largest root of $\mu[v_1, \ldots, v_m](x)$ is at most
\[
t + \delta = 2\sqrt{2\epsilon + \epsilon^2}.
\]

\[\square\]

**Proof of Theorem 1.2.** Let $\epsilon = \epsilon_1 + \epsilon_2$ as always. Theorem 4.1 implies that the largest root of the mixed characteristic polynomial, $\mu[v_1, \ldots, v_m](x)$, is at most $2\sqrt{2\epsilon + \epsilon^2}$. Theorem 3.3 tells us that the polynomials $\{q_S\}_{S \in \mu(S) > 0}$ form an interlacing family. So, by Theorem 2.4 there is a set $S \subseteq [m]$ with $\mu(S) > 0$ such that the largest root of
\[
\det \left( x^2 I - \sum_{i \in S} 2v_i v_i^\top \right)
\]
is at most $2\sqrt{2\epsilon + \epsilon^2}$. This implies that the largest root of
\[
\det \left( x I - \sum_{i \in S} 2v_i v_i^\top \right)
\]
is at most $(2\sqrt{2\epsilon + \epsilon^2})^2$. Therefore,
\[
\left\| \sum_{i \in S} v_i v_i^\top \right\| = \frac{1}{2} \left\| \sum_{i \in S} 2v_i v_i^\top \right\| \leq \frac{1}{2} (2\sqrt{2\epsilon + \epsilon^2})^2 = 4\epsilon + 2\epsilon^2.
\]

\[\square\]
5 Discussion

Similar to [MSS13b] our main theorem is not algorithmic, i.e., we are not aware of any polynomial time algorithm that for a given homogeneous strongly Rayleigh distribution with small marginal probabilities and for a set of vectors assigned to the underlying elements with small norm finds a sample of the distribution with spectral norm bounded away from 1. Such an algorithm can lead to improved approximation algorithms for the Asymmetric Traveling Salesman Problem.

Although our main theorem can be seen as a generalization of [MSS13b] the bound that we prove on the maximum root of the mixed characteristic polynomial is incomparable to the bound of Theorem 1.1. We can use our main theorem to prove Weaver’s $K_{2r}$ conjecture [Wea04] for $r > 4$. It is an interesting question to see if the dependency on $\epsilon$ in our multivariate barrier can be improved.

Acknowledgement

We would like to thank Adam Marcus, Dan Spielman, and Nikhil Srivastava for stimulating discussions regarding the main obstacles in generalizing their proof of the Kadison-Singer problem.

References

[AA91] Charles A. Akemann and Joel Anderson. Lyapunov theorems for operator algebras, volume 94. Memoirs of the American Mathematical Society, 1991.

[AGM+10] Arash Asadpour, Michel X. Goemans, Aleksander Madry, Shayan Oveis Gharan, and Amin Saberi. An $o(\log n/\log \log n)$-approximation algorithm for the asymmetric traveling salesman problem. In SODA, pages 379–389, 2010.

[AO14] Nima Anari and Shayan Oveis Gharan. Effective-resistance-reducing flows and asymmetric tsp. 2014.

[BB08] Julius Borcea and Petter Brändén. Applications of stable polynomials to mixed determinants: Johnson’s conjectures, unimodality, and symmetrized fischer products. Duke Math. Journal, 143(2):205–223, 2008.

[BB10] J. Borcea and P. Brändén. Multivariate pólya-schur classification problems in the weyl algebra. Proceedings of the London Mathematical Society, 101(3):73–104, 2010.

[BBL09] Julius Borcea, Petter Branden, and Thomas M. Liggett. Negative dependence and the geometry of polynomials. Journal of American Mathematical Society, 22:521–567, 2009.

[BSS14] Joshua D. Batson, Daniel A. Spielman, and Nikhil Srivastava. Twice-ramanujan sparsifiers. SIAM Review, 56(2):315–334, 2014.

[Ded92] Jean Pierre Dedieu. Obreschkoff’s theorem revisited: what convex sets are contained in the set of hyperbolic polynomials? Journal of Pure and Applied Algebra, 81(3):269–278, 1992.

[HO14] Nicholas J. A. Harvey and Neil Olver. Pipage rounding, pessimistic estimators and matrix concentration. In SODA, pages 926–945, 2014.
Richard V. Kadison and I. M. Singer. Extensions of pure states. *American Journal of Mathematics*, 81(2):383–400, 1959.

A. Lewis, P. Parrilo, and M. Ramana. The lax conjecture is true. *Proceedings of the American Mathematical Society*, 133(9), 2005.

Adam Marcus, Daniel A. Spielman, and Nikhil Srivastava. Interlacing families I: bipartite ramanujan graphs of all degrees. In *FOCS*, pages 529–537, 2013.

Adam Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families ii: Mixed characteristic polynomials and the kadison-singer problem. 2013.

C. St. J. A. Nash-Williams. Edge disjoint spanning trees of finite graphs. *J. London Math. Soc.*, 36:445–45, 1961.

Shayan Oveis Gharan, Amin Saberi, and Mohit Singh. A randomized rounding approach to the traveling salesman problem. In *FOCS*, pages 550–559, 2011.

Robin Pemantle and Yuval Peres. Concentration of lipschitz functionals of determinantal and other strong rayleigh measures. *Combinatorics, Probability and Computing*, 23:140–160, 2014.

Mohit Singh and Nisheeth K. Vishnoi. Entropy, optimization and counting. In *STOC*, pages 50–59, 2014.

David G. Wagner. Multivariate stable polynomials: theory and applications. *Bulletin of the American Mathematical Society*, 48(1):53–53, January 2011.

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