IMPROVED BOUNDS IN WEAVER AND FEICHTINGER CONJECTURES

MARCIN BOWNIK, PETE CASAZZA, ADAM W. MARCUS, AND DARRIN SPEEGLE

Abstract. We sharpen the constant in the KS_2 conjecture of Weaver [31] that was given by Marcus, Spielman, and Srivastava [28] in their solution of the Kadison–Singer problem. We then apply this result to prove optimal asymptotic bounds on the size of partitions in the Feichtinger conjecture.

1. Introduction

The goal of this paper is to explore some consequences of the recent resolution [28] of the Kadison–Singer problem [26]. The Kadison–Singer problem was known to be equivalent to a large number of problems in analysis such as the Anderson Paving Conjecture [2, 3, 4], Bourgain–Tzafriri Restricted Invertibility Conjecture [9, 10, 11], Akemann–Anderson Projection Paving Conjecture [1], Feichtinger Conjecture [14, 15, 22], R_ϵ Conjecture [19], and Weaver Conjecture [31]. For an extensive study of problems equivalent to the Kadison–Singer problem we refer to [16, 18, 19]. Consequently, the breakthrough resolution of the Weaver Conjecture [31] by Marcus, Spielman, and Srivastava [28] automatically validates all of these conjectures. At the same time, it raises the question of finding optimal quantitive bounds in these problems.

In this paper we shall concentrate on showing quantitative bounds in Weaver and Feichtinger Conjectures. The first part of the paper focuses on improving bounds to the conjecture of Weaver known as KS_2. The proof of the KS_2 conjecture relies on the following result [28, Theorem 1.4].

Theorem 1.1 (Marcus–Spielman–Srivastava)**.** *If $\epsilon > 0$ and $v_1, \dots, v_m$ are independent random vectors in $\mathbb{C}^d$ with finite support such that*

$$\sum_{i=1}^m \mathbb{E}\left[v_i v_i^*\right] = \mathbf{I} \quad \text{and} \quad \mathbb{E}\left[\|v_i\|^2\right] \le \epsilon$$

for all i, then

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

We show the following sharpening of Theorem 1.1.

Date: May 17, 2018.

The first author was partially supported by NSF grant DMS-1265711. The second author was supported by NSF DMS 1307685; NSF ATD 1321779; and AFOSR: FA9550-11-1-0245. The fourth author was partially supported by a grant from the Simons Foundation #244953. The work on this paper was initiated during the AIM workshop "Beyond Kadison–Singer: paving and consequences" on December 1-5, 2014.

Theorem 1.2. *If $0 < \epsilon < 1/2$ and $v_1, \dots, v_m$ are independent random vectors in $\mathbb{C}^d$ with support of size 2 such that*

$$\sum_{i=1}^{m} \mathbb{E}\left[v_i v_i^*\right] = \mathbf{I} \quad \text{and} \quad \mathbb{E}\left[\|v_i\|^2\right] \leq \epsilon \tag{1.1}$$

for all i, then

$$\mathbb{P}\left(\left\|\sum_{i=1}^{m} v_i v_i^*\right\| \leq 1 + 2\sqrt{\epsilon}\sqrt{1-\epsilon}\right) > 0. \tag{1.2}$$

Theorem 1.2 leads to improved bounds in the conjecture of Weaver known as KS_2. Corollary 1.3 improves the original methods of [28] that yield the same result albeit for constants $\eta > (2+\sqrt{2})^2 \approx 11.6569$.

Corollary 1.3. *For every $\eta > 4$, there exist $\theta > 0$ such that the following holds. Let $u_1, \dots, u_m \in \mathbb{C}^d$ be such that $\|u_i\| \leq 1$ for all i and*

$$\sum_{i=1}^{m} |\langle u, u_i\rangle|^2 = \eta \qquad \text{for all } ||u|| = 1. \tag{1.3}$$

Then there exists a partition of $[m] := \{1, \dots, m\}$ into sets I_1 and I_2 so that for $k = 1, 2$,

$$\sum_{i \in I_k} |\langle u, u_i\rangle|^2 \leq \eta - \theta \qquad \text{for all } ||u|| = 1. \tag{1.4}$$

In the second part of the paper we shall deduce quantitative bounds for the Feichtinger conjecture. As a consequence of Corollary 1.3 we show that any Parseval frame $\{v_i\}_{i\in I} \subset \mathcal{H}$ (or more generally a Bessel sequence with bound 1) with norms $\|v_i\| \geq \varepsilon$, where $\varepsilon > \sqrt{3}/2$, can be decomposed into two Riesz sequences. We also show the following asymptotic estimate on the size of the partition as ε approaches to 0.

Theorem 1.4. *Suppose $\{v_i\}_{i\in I}$ is a Bessel sequence for a separable Hilbert space $\mathcal{H}$ with bound 1 that consists of vectors of norms $\|v_i\| \geq \varepsilon$, where $\varepsilon > 0$. Then there exists a universal constant $C > 0$, such that I can be partitioned into $r \leq C/\varepsilon^2$ subsets $I_1, \dots, I_r$, such that each subfamily $\{v_i\}_{i\in I_j}$, $j = 1, \dots, r$, is a Riesz sequence.*

It is easy to see that Theorem 1.4 gives the optimal asymptotic behavior on the size r of the partition. Indeed, it suffices to consider the union of $r = \lfloor 1/\varepsilon^2 \rfloor$ copies of an orthogonal basis of $\mathcal{H}$ scaled by a factor ε. This yields a Bessel sequence with bound 1 that can not be partitioned into fewer than r Riesz sequences.

1.1. **Review of Marcus, Spielman, Srivastava.** We give a brief review of the proof in [28] to provide some context for the statement of our main technical theorem. See Section 4 for a more detailed discussion.

The results in [28] use a construction introduced by the same authors in [27] that they called an *interlacing family* of polynomials. In [27], they showed that each interlacing family constructed from a collection of polynomials $\{p_i\}$ provides a polynomial p^* with the following properties:

(1) p^* has all real roots,
(2) $\text{maxroot}(p^*) > \text{maxroot}(p_j)$ for some j.

As a result, if one can bound the largest root of the associated p^*, then one can assert that some polynomial in the collection has a largest root which satisfies the same bound.

In order to apply this to matrices, [28] uses characteristic polynomials. They consider certain convex combinations of these polynomials, which they call *mixed characteristic polynomials*. To bound the largest root of the mixed characteristic polynomial, they define a process on multivariate polynomials which starts at a determinantal polynomial and ends at a polynomial which is a multivariate version of the mixed characteristic polynomial. They use what they call a *barrier function* to maintain an upper bound on the size of the largest root as the process evolves.

The bound that was proved in [28] holds for mixed characteristic polynomials in general. In the first part of this paper, we consider the special case of when they are (at most) quadratic in each of its variables (corresponding to matrices of rank at most 2). Our main technical theorem is the following:

Theorem 1.5. *Suppose $A_1, \dots, A_m$ are $d \times d$ Hermitian positive semidefinite matrices of rank at most 2 satisfying*

$$\sum_{i=1}^{m} A_i = \mathbf{I} \quad \textit{and} \quad \operatorname{Tr}(A_i) \leq \epsilon < 1/2 \quad \textit{for all } i.$$

Then the largest root of the polynomial

$$\left(\prod_{i=1}^{m} 1 - \partial_{y_i} \right) \det \left[\sum_{i=1}^{m} y_i A_i \right] \Big|_{y_1 = \cdots = y_m = x} \tag{1.5}$$

is at most $1 + 2\sqrt{\epsilon}\sqrt{1-\epsilon}$.

This is an improvement over the value of $(1+\sqrt{\epsilon})^2$ in [28, Theorem 5.1], but only in the case of rank 2 matrices. The proof follows the general outline in [28], but employs tighter inequalities that exploit the bounded rank of the matrices in (1.5). Our main analytic tool will be the *mixed discriminant*, a multilinear generalization of the determinant function. In Section 2 we will review the properties of the mixed discriminant that we will need in later sections. Some of these properties are well known (see, for example [7, 23]), but for the benefit of the reader we will try to make the presentation self-contained.

1.2. **Organization.** The paper is organized as follows. In Section 2 we present some elementary properties of mixed discriminants and then in Section 3 we show how these properties can be used to establish bounds on the barrier function discussed in the previous section. Section 4, in particular, contains the proofs of Theorem 1.2, Corollary 1.3, and Theorem 1.5.

In Sections 5 and 6, we apply Theorems 1.1 and 1.2 to get quantitative bounds in frame theory. Section 5 contains the results from frame theory that show the interlinking properties of complementary subsets of a Parseval frame. In Section 6 we use the results in the previous sections to explore implications in frame theory. In particular, we prove Theorem 1.4 and some of its variants such as the R_ϵ conjecture and Bourgain–Tzafriri conjecture. Our focus here will be in optimizing the bounds that follow from the results of the previous sections.

2. Mixed discriminant and properties

Let S_d denote the symmetric group on d elements. Given $d \times d$ matrices $X_1, \ldots, X_d$ and a permutation $\sigma \in S_d$, let $Y_\sigma(X_1, \ldots, X_d)$ be the matrix with

$$\mathrm{col}_j \left[Y_\sigma(X_1, \ldots, X_n)\right] = \mathrm{col}_j \left[X_{\sigma(j)}\right]$$

where col_j denotes the "jth column" function. Y_σ can be seen as a "mixture" of its input matrices since each of its columns comes from a different input matrix.

Definition 2.1. The *mixed discriminant* of $X_1, \ldots, X_d$ is the quantity

$$D\left(X_1, \ldots, X_d\right) = \sum_{\sigma \in S_d} \det\left[Y_\sigma(X_1, \ldots, X_d)\right].$$

Remark 2.2. Note that our definition of the mixed discriminant differs by a factor of $d!$ from many other treatments (including [7]), corresponding to an average over S_d rather than a sum. The literature is far from standard in this respect, and our reason for taking this normalization is that it will simplify a number of the formulas we will use.

To ease notation slightly, given a matrix X, we will write $X[k]$ to denote a vector that repeats X k times. The following two examples follow directly from Definition 2.1.

Example 2.3. For X a $d \times d$ matrix,

$$D\left(X[d]\right) = D\left(X, X, \ldots, X\right) = d! \det\left[X\right].$$

Example 2.4. For any $d \times d$ matrix X,

$$D\left(X, \mathbf{I}[d-1]\right) = (d-1)! \,\mathrm{Tr}(X).$$

It should be clear that $D\left(\right)$ is symmetric in its arguments (we will refer to this property as *permutation invariance*). One property of the mixed discriminant that gives it much of its versatility as an analytic tool is its multilinearity (linearity in each input matrix). Due to permutation invariance, it suffices to state this as linearity in the first coordinate.

Lemma 2.5 (Multilinearity)**.**

$$D\left(aA + bB, X_2, \ldots, X_d\right) = aD\left(A, X_2, \ldots, X_d\right) + bD\left(B, X_2, \ldots, X_d\right)$$

Proof. It suffices to show that

$$\det\left[Y_\sigma(aA + bB, X_2, \ldots, X_d)\right] = a \det\left[Y_\sigma(A, X_2, \ldots, X_d)\right] + b \det\left[Y_\sigma(B, X_2, \ldots, X_d)\right]$$

for each permutation σ as then the same holds for the sum. However, this follows easily from the definition and the linearity of the determinant with respect to columns:

$$\det\left[\; a\vec{u} + b\vec{v} \mid \vec{x}_2 \mid \ldots \mid \vec{x}_d \;\right] = a \det\left[\; \vec{u} \mid \vec{x}_2 \mid \ldots \mid \vec{x}_d \;\right] + b \det\left[\; \vec{v} \mid \vec{x}_2 \mid \ldots \mid \vec{x}_d \;\right]$$

since exactly one column of Y_σ comes from any one of its inputs. □

One useful corollary of multilinearity is that the mixed discriminant has an expansion similar to the binary expansion

$$(a+b)^d = \sum_i \binom{d}{i} a^i b^{d-i}.$$

Starting with Example 2.3 and iterating Lemma 2.5 gives an analogous formula:

Example 2.6.

$$d!\det[xA+yB] = D((xA+yB)[d]) = \sum_i \binom{d}{i} x^i y^{d-i} D(A[i], B[d-i]).$$

Mixed discriminants also have useful multiplicative properties, which are not as easily inferred from Definition 2.1. For this reason, we find it worthwhile to derive an equivalent characterization (which one often sees given as the primary definition).

Lemma 2.7. *For $d \times d$ matrices $X_1, \ldots, X_d$, we have*

$$D(X_1, \ldots, X_d) = \frac{\partial^d}{\partial t_1 \ldots \partial t_d} \det\left[\sum_{i=1}^d t_i X_i\right].$$

Proof. Note that $\det\left[\sum_{i=1}^d t_i X_i\right]$ is a homogeneous degree d polynomial. By the linearity of the determinant with respect to columns we have

$$\det\left[\sum_{i=1}^d t_i X_i\right] = \sum_{i_1=1}^d \ldots \sum_{i_d=1}^d t_{i_1} \cdots t_{i_d} \det[\mathrm{col}_1[X_{i_1}] | \ldots | \mathrm{col}_d[X_{i_d}]].$$

Since

$$\frac{\partial^d (t_{i_1} \cdots t_{i_d})}{\partial t_1 \ldots \partial t_d} = \begin{cases} 1 & \text{if } (i_1, \ldots, i_d) \text{ is a permutation of } [d], \\ 0 & \text{otherwise,} \end{cases}$$

the partial derivative $\frac{\partial^d}{\partial t_1 \ldots \partial t_d}$ will pick up only terms corresponding to a permutation of $[d]$. Hence,

$$\frac{\partial^d}{\partial t_1 \ldots \partial t_d} \det\left[\sum_{i=1}^d t_i X_i\right] = \sum_{\sigma \in S_d} \det\left[\mathrm{col}_1\left[X_{\sigma(1)}\right] | \ldots | \mathrm{col}_d\left[X_{\sigma(d)}\right]\right] = D(X_1, \ldots, X_d)$$

as required. □

The characterization in Lemma 2.7 is often easier to work with than Definition 2.1. This is evident in the following example:

Example 2.8. Let $X_1, \ldots, X_d$ be $d \times d$ matrices and $\vec{v}$ a vector such that $X_k \vec{v} = \vec{0}$ for all k. Then

$$D(X_1, \ldots, X_d) = 0$$

Expressing higher rank matrices as sums of rank 1 matrices can often simplify proofs considerably. In such cases, the following lemma is quite useful:

Lemma 2.9. *If $d \times d$ matrices $X_1, \ldots, X_d$ have rank 1, then*

$$D(X_1, \ldots, X_d) = \det\left[\sum_{i=1}^d X_i\right].$$

Proof. Let $X_i = u_i v_i^*$. We first note that if $v_i = v_j$ or $u_i = u_j$ for any $i \neq j$, then $\sum_{i=1}^d t_i u_i v_i^*$ would have rank less than d. Hence by Example 2.8, we have

$$D(X_1, \ldots, X_d) = \det\left[\sum_{i=1}^d t_i X_i\right] = 0.$$

So assume that vectors u_i (and, separately, vectors v_i) are distinct. By Example 2.3, we have

$$\det\left[\sum_{i=1}^{d} u_i v_i^*\right] = \frac{1}{d!} D\left(\left(\sum_{i=1}^{d} u_i v_i^*\right)[d]\right) = \sum_{i_1=1}^{d} \cdots \sum_{i_d=1}^{d} D\left(u_{i_1} v_{i_1}^*, \ldots, u_{i_d} v_{i_d}^*\right)$$

where any term with two indices $i_j = i_k$ contributes 0 (for the reason mentioned above). Hence the only contributing terms happen if $(i_1, \ldots, i_d)$ is a permutation of $[d]$ and there are $d!$ of these. Furthermore, each of these gives the same contribution, since $D\,()$ is symmetric in its arguments. Hence, we have

$$\det\left[\sum_{i} u_i v_i^*\right] = \frac{1}{d!} D\left(\left(\sum_{i} u_i v_i^*\right)[d]\right) = D\left(u_1 v_1^*, \ldots, u_d v_d^*\right).$$

□

The next lemma is an extension of the familiar multiplication identity of the determinant $\det[AB] = \det[A]\det[B]$.

Lemma 2.10 (Multiplication)**.** *For $d \times d$ matrices $X_1, \ldots, X_d, Y_1, \ldots, Y_d$, we have*

$$D\left(X_1, \ldots, X_d\right) D\left(Y_1, \ldots, Y_d\right) = \sum_{\pi \in S_d} D\left(X_1 Y_{\pi(1)}, \ldots, X_d Y_{\pi(d)}\right)$$

Proof. Using the characterization in Lemma 2.7, we can write

$$\begin{aligned} D\left(X_1, \ldots, X_d\right) D\left(Y_1, \ldots, Y_d\right) &= \frac{\partial^d}{(\partial t_1)\ldots(\partial t_d)} \frac{\partial^d}{(\partial s_1)\ldots(\partial s_d)} \det\left[\sum_{i=1}^{d} t_i X_i\right] \det\left[\sum_{i=1}^{d} s_i Y_i\right] \\ &= \frac{\partial^d}{(\partial t_1)\ldots(\partial t_d)} \frac{\partial^d}{(\partial s_1)\ldots(\partial s_d)} \det\left[\sum_{i=1}^{d}\sum_{j=1}^{d} t_i s_j X_i Y_j\right]. \end{aligned}$$

Expanding the determinant using multilinearity will result in a homogeneous polynomial of degree $2d$ in the variables $s_1, \ldots, s_d, t_1, \ldots, t_d$, where each term will be of the form

$$D\left(X_{i_1} Y_{j_1}, \ldots, X_{i_d} Y_{j_d}\right) s_{i_1} \ldots s_{i_d} t_{j_1} \ldots t_{j_d}.$$

The lemma then follows by noticing that the coefficients $D\left(X_1 Y_{\pi(1)}, \ldots, X_d Y_{\pi(d)}\right)$ for some permutation π are exactly the ones that will remain after the differentiations. □

The following examples are basic applications of Lemma 2.10:

Example 2.11.

$$\begin{aligned} d!\det[A]\, D\left(X_1, \ldots, X_d\right) &= D\left(A[d]\right) D\left(X_1, \ldots, X_d\right) \\ &= d! D\left(AX_1, \ldots, AX_d\right) = d! D\left(X_1 A, \ldots, X_d A\right). \end{aligned}$$

Example 2.12.

$$\begin{aligned} D\left(A, \mathbf{I}[d-1]\right) D\left(X_1, X_2, \ldots, X_d\right) = (d-1)!\Big(& D\left(AX_1, X_2, \ldots, X_d\right) \\ & + D\left(X_1, AX_2, \ldots, X_d\right) + \cdots + D\left(X_1, X_2, \ldots, AX_d\right)\Big) \end{aligned}$$

We now extend Definition 2.1 slightly so as to ease notation even further.

Definition 2.13. For $d \times d$ matrices $X_1, \ldots, X_k$ where $k \leq d$, we will write

$$\tilde{D}(X_1, \ldots, X_k) = \frac{D(X_1, \ldots, X_k, \mathbf{I}[d-k])}{(d-k)!}$$

where $\mathbf{I}$ is the $d \times d$ identity matrix.

In particular, note that $D()$ and $\tilde{D}()$ are equivalent when there are d matrices and that $\tilde{D}(\varnothing) = 1$. Using Example 2.11 and Example 2.12 above, we get the following two corollaries:

Corollary 2.14. *Let $A, X_1, \ldots, X_k$ be $d \times d$ matrices with $k \leq d$ and with A positive semi-definite. Then*

$$\tilde{D}(AX_1, \ldots, AX_k) = \tilde{D}\left(A^{1/2}X_1A^{1/2}, \ldots, A^{1/2}X_kA^{1/2}\right).$$

Proof.

$$\begin{aligned}
\tilde{D}(AX_1, \ldots, AX_k) &= \frac{1}{(d-k)!} D(AX_1, \ldots, AX_k, \mathbf{I}[d-k]) \\
&= \frac{1}{(d-k)!} \det\left[A^{1/2}\right] D\left(A^{1/2}X_1, \ldots, A^{1/2}X_k, A^{-1/2}\mathbf{I}[d-k]\right) \\
&= \frac{1}{(d-k)!} D\left(A^{1/2}X_1A^{1/2}, \ldots, A^{1/2}X_kA^{1/2}, A^{-1/2}\mathbf{I}A^{1/2}[d-k]\right) \\
&= \tilde{D}\left(A^{1/2}X_1A^{1/2}, \ldots, A^{1/2}X_kA^{1/2}\right)
\end{aligned}$$

□

Corollary 2.15. *If $A, X_1, \ldots, X_k$ are $d \times d$ matrices with $k \leq d$, then*

$$\tilde{D}(A)\,\tilde{D}(X[k]) = k\tilde{D}(AX, X[k-1]) + \tilde{D}(A, X[k]).$$

Proof. By definition,

$$\tilde{D}(A)\,\tilde{D}(X[k]) = \frac{1}{(d-k)!}\frac{1}{(d-1)!} D(A, \mathbf{I}[d-1])\, D(X[k], \mathbf{I}[d-k])$$

and note that there are $k(d-1)!$ permutations π where $\pi(1) \leq k$ and $(d-k)(d-1)!$ others. So by Lemma 2.10, we have

$$\begin{aligned}
D(A, \mathbf{I}[d-1])\, D(X[k], \mathbf{I}[d-k]) = {} & k(d-1)!D(AX, X[k-1], \mathbf{I}[d-k]) \\
& + (d-k)(d-1)!D(A, X[k], \mathbf{I}[d-k-1])
\end{aligned}$$

and combining the two gives

$$\begin{aligned}
\tilde{D}(A)\,\tilde{D}(X[k]) &= \frac{k}{(d-k)!} D(AX, X[k-1], \mathbf{I}[d-k]) \\
&\quad + \frac{1}{(d-k-1)!} D(A, X[k], \mathbf{I}[d-k-1]) \\
&= k\tilde{D}(AX, X[k-1]) + \tilde{D}(A, X[k])
\end{aligned}$$

as claimed. □

Lastly, we will require two straightforward inequalities:

Lemma 2.16 (Positivity)**.** *If $X_1, \ldots, X_k$ are $d \times d$ positive semidefinite matrices with $k \leq d$, then*

$$\tilde{D}(X_1, \ldots, X_k) \geq 0.$$

Proof. Note that it suffices to prove this when $k = d$ since

$$\tilde{D}(X_1, \ldots, X_k) = \frac{1}{(d-k)!} D(X_1, \ldots, X_k, \mathbf{I}[d-k])$$

and so the left hand side is nonnegative exactly when the right hand side is. However the case when $X_1, \ldots, X_d$ have rank 1 follows from Lemma 2.9, and then the general case follows by multilinearity. □

The following fact is a special case of a result due to Artstein-Avidan, Florentin, and Ostrover [6, Theorem 1.1].

Lemma 2.17. *If A and B are $d \times d$ positive semidefinite matrices with $d \geq 2$, then*

$$\tilde{D}(A)\,\tilde{D}(B) \geq \tilde{D}(A, B).$$

Proof. By Lemma 2.10, we have

$$\begin{aligned}
((d-1)!)^2 \tilde{D}(A)\,\tilde{D}(B) &= D(A, \mathbf{I}[d-1])\, D(B, \mathbf{I}[d-1]) \\
&= (d-1)! D(AB, \mathbf{I}[d-1]) + (d! - (d-1)!) D(A, B, \mathbf{I}[d-2]) \\
&= ((d-1)!)^2 \operatorname{Tr}(AB) + (d-1)(d-1)! D(A, B, \mathbf{I}[d-2]) \\
&= ((d-1)!)^2 \operatorname{Tr}(AB) + ((d-1)!)^2 \tilde{D}(A, B).
\end{aligned}$$

Rearranging gives

$$\tilde{D}(A)\,\tilde{D}(B) - \tilde{D}(A, B) = \operatorname{Tr}(AB)$$

which is nonnegative whenever A and B are positive semidefinite. □

Finally, we would like to mention a recent characterization of mixed discriminants by Florentin, Milman, and Schneider [21, Theorem 2]. Up to a multiplicative constant, the mixed discriminant is the unique function on d-tuples of positive semidefinite matrices that is multilinear, non-negative, and which is zero if two of its arguments are proportional matrices of rank one.

3. Application to polynomials

Let $\mathbb{H} \subset \mathbb{C} = \{z : \Im(z) > 0\}$ (where $\Im$ denotes the "imaginary part"). A polynomial $p \in \mathbb{C}[x_1, \ldots, x_m]$ is called *stable* if $\vec{y} \in \mathbb{H}^m$ implies $p(\vec{y}) \neq 0$. A polynomial is called *real stable* if it is stable and all of its coefficients are real.

The connection between mixed discriminants and real stable polynomials can be derived from an incredibly useful result of Helton and Vinnikov [24]. Here, we will use an extension that specializes to our case of interest [8, Corollary 6.7]:

Theorem 3.1. *Let $p(x, y)$ be a degree d real stable polynomial. Then there exist $d \times d$ real symmetric matrices A, B, C such that*

$$p(x, y) = \pm \det[xA + yB + C].$$

Furthermore, A and B can be taken to be positive semidefinite.

Remark 3.2. We should note that the representation via real symmetric matrices in Theorem 3.1 is actually quite a bit stronger than is needed for the results in this paper. For our purposes, it would suffice to have a representation using Hermitian matrices (a far weaker constraint, both theoretically and computationally, see [30]).

We would like to understand the behavior of a given real stable polynomial at a selected reference point $\vec{z}$. Recall the following definition from [28]:

Definition 3.3. Let $Q(x_1, \ldots, x_m)$ be a multivariate polynomial. We say that a reference point $\vec{z} \in \mathbb{R}^m$ is *above the roots* of Q if

$$Q(\vec{z} + \vec{s}) > 0 \qquad \text{for all} \qquad \vec{s} = (s_1, \ldots, s_m) \in \mathbb{R}^m, s_i \geq 0,$$

i.e., if Q is positive on the nonnegative orthant with origin at $\vec{z}$.

In the case that a reference point is above the roots of a polynomial, a more specific version of Theorem 3.1 can be obtained.

Corollary 3.4. *Let $p(x, y)$ be a degree d real stable polynomial with (x_0, y_0) above the roots of p. Then there exist $d \times d$ real symmetric matrices A, B, C such that*

$$p(x, y) = \det [xA + yB + C]$$

and the following hold:

- *A and B are positive semidefinite*
- *$A + B$ is positive definite*
- *$M = x_0A + y_0B + C$ is positive definite*

Proof. Let

$$p(x, y) = \pm \det [xA + yB + C]$$

be the representation provided by Theorem 3.1. Now let

$$q(t) = p(t, t) = \pm \det [t(A + B) + C] .$$

Since p has (total) degree d, q must have degree d (in t) and so we must have $\det [A + B] \neq 0$. Given that A and B are each positive semidefinite, this ensures $A + B$ is positive definite. Furthermore, since (x_0, y_0) is above the roots of p, we have $q(t) > 0$ for all for $t \geq \max\{x_0, y_0\}$. Hence q must have a positive first coefficient, which means

$$p(x, y) = \det [xA + yB + C] .$$

Thus it remains to show that $M = x_0A + y_0B + C$ is positive definite. To see this, consider the matrices

$$M_t = (x_0 + t)A + (y_0 + t)B + C.$$

Since $A + B$ is positive definite, M_t is positive definite for large enough t. Now note that for $t \geq 0$, we have

$$\det [M_t] = p(x_0 + t, y_0 + t) > 0$$

since (x_0, y_0) is above the roots of p. This implies that the minimum eigenvalue of M_t (which is a continuous function in t) remains above 0 for all $t \geq 0$, and so (in particular) $M = M_0$ is positive definite. □

For the remainder of the section, we will fix a degree d real stable polynomial $p(x, y)$ and a reference point (x_0, y_0) above the roots of p. We also fix the matrices A, B, C provided by Corollary 3.4 and set $M = x_0A + y_0B + C$. Since M is positive definite, it has a well defined square root, and so we can define the matrices

$$\hat{A} = M^{-1/2}AM^{-1/2} \quad \text{and} \quad \hat{B} = M^{-1/2}BM^{-1/2}.$$

Note that $\hat{A}$ and $\hat{B}$ are both positive semidefinite, since A, B and M are.

Corollary 3.5. *The matrices $\hat{A}$ and $\hat{B}$ defined above satisfy*

$$\frac{\partial^i}{(\partial x)^i}\frac{\partial^j}{(\partial y)^j}p(x_0, y_0) = \tilde{D}\left(\hat{A}[i], \hat{B}[j]\right)p(x_0, y_0)$$

for all $i + j \leq d$.

Proof. Let

$$p(x_0 + \epsilon, y_0 + \delta) = \sum_{i=0}^{d}\sum_{j=0}^{d-i}\frac{\epsilon^i}{i!}\frac{\delta^j}{j!}c_{i,j}$$

be the Taylor expansion of a polynomial p of degree d. Our goal is to show that $c_{i,j} = \tilde{D}\left(\hat{A}[i], \hat{B}[j]\right)p(x_0, y_0)$. By Example 2.3, we can write

$$p(x_0 + \epsilon, y_0 + \delta) = \det\left[(x_0 + \epsilon)A + (y_0 + \delta)B + C\right] = \frac{1}{d!}D\left((M + \epsilon A + \delta B)[d]\right)$$

and using multilinearity, we have

$$p(x_0 + \epsilon, y_0 + \delta) = \frac{1}{d!}\sum_{i=0}^{d}\sum_{j=0}^{d-i}\epsilon^i\delta^j\binom{d}{i, j, d-i-j}D\left(A[i], B[j], M[d-i-j]\right),$$

where $\binom{d}{i,j,d-i-j}$ is the multinomial coefficient $\frac{d!}{i!j!(d-i-j)!}$. Equating coefficients then gives

$$\begin{aligned} c_{i,j} &= \frac{i!j!}{d!}\binom{d}{i, j, d-i-j}D\left(A[i], B[j], M[d-i-j]\right) \\ &= \frac{1}{(d-i-j)!}D\left(A[i], B[j], M[d-i-j]\right). \end{aligned}$$

By Corollary 2.14, we can factor out the M term to get

$$c_{i,j} = \frac{1}{(d-i-j)!}\det[M]\, D\left(\hat{A}[i], \hat{B}[j], \mathbf{I}[d-i-j]\right) = \det[M]\,\tilde{D}\left(\hat{A}[i], \hat{B}[j]\right)$$

which, since $\det[M] = p(x_0, y_0)$, is exactly the claimed result. □

Corollary 3.5 provides a way to associate the partial derivatives of a bivariate real stable polynomial at a point to a mixed discriminant involving the matrices in its determinantal representation. This, coupled with Corollary 3.4, will allow us to use properties of positive semidefinite matrices to derive inequalities for points above the roots of p. With this in mind, we will attempt to quantify the concept of being "above the roots". For a polynomial q, we will consider the barrier function

$$\Phi_q^x = \begin{cases} \frac{\partial}{\partial x}\ln q(x_0, y_0) & \text{when } (x_0, y_0) \text{ is above the roots of } q \\ \infty & \text{otherwise} \end{cases}$$

and Φ_q^y defined similarly (except with the derivative in the y coordinate). In particular, we will be interested in the behavior of these functions under transformations of q. The next lemma is a general result in that direction:

Lemma 3.6. *Let $R = \sum_i a_i \partial_x^i$ be a differential operator with real coefficients $\{a_i\}$ such that $q = R(p)$ and such that (x_0, y_0) is above the roots of both p and q. Then the following two statements are equivalent:*

(1) $\Phi_q^y \leq \Phi_p^y$

(2) $\sum_i i a_i \tilde{D}\left(\hat{A}[i-1], \hat{L}\right) \geq 0$, where $\hat{L} = \hat{A}^{1/2}\hat{B}\hat{A}^{1/2}$ is positive semidefinite.

Proof. Using Corollary 3.5, we can write

$$\frac{q}{p} = \sum_i a_i \tilde{D}\left(\hat{A}[i]\right)$$

so that

$$\frac{q_y}{p} = \sum_i a_i \tilde{D}\left(\hat{A}[i], \hat{B}\right) \quad \text{and} \quad \frac{qp_y}{p^2} = \sum_i a_i \tilde{D}\left(\hat{A}[i]\right) \tilde{D}\left(\hat{B}\right).$$

Now using Corollary 2.15, we have

$$\tilde{D}\left(\hat{A}[i]\right) D\left(\hat{B}\right) = \tilde{D}\left(\hat{A}[i], \hat{B}\right) + i\tilde{D}\left(\hat{A}[i-1], \hat{A}\hat{B}\right).$$

By Corollary 3.4, both $\hat{A}$ and $\hat{B}$ are positive semidefinite, and so we can apply Corollary 2.14 to get

$$\tilde{D}\left(\hat{A}[i], \hat{A}\hat{B}\right) = \tilde{D}\left(\hat{A}[i], \hat{L}\right)$$

where $\hat{L} = \hat{A}^{1/2}\hat{B}\hat{A}^{1/2}$ is positive semidefinite. Combining these gives

$$\frac{pq_y - qp_y}{p^2} = \sum_i a_i \left(\tilde{D}\left(\hat{A}[i], \hat{B}\right) - \tilde{D}\left(\hat{A}[i]\right) \tilde{D}\left(\hat{B}\right)\right) = -\sum_i i a_i \tilde{D}\left(\hat{A}[i-1], \hat{L}\right).$$

Therefore we can write

$$\Phi_q^y - \Phi_p^y = \frac{pq_y - qp_y}{pq} = \left(\frac{p}{q}\right)\frac{pq_y - qp_y}{p^2} = -\left(\frac{p}{q}\right)\sum_i i a_i \tilde{D}\left(\hat{A}[i-1], \hat{L}\right)$$

where p/q is positive when (x_0, y_0) is above the roots of both p and q. Hence

$$\Phi_q^y - \Phi_p^y \quad \text{and} \quad -\sum_i i a_i \tilde{D}\left(\hat{A}[i-1], \hat{L}\right)$$

have the same sign, as required. □

Using this machinery, we now prove two lemmas that will help us exploit the quadratic nature of the polynomials. The first of these lemmas is a strengthening of [28, Lemma 5.11] in the case that the polynomial is quadratic.

Lemma 3.7. *Assume $p(x, y)$ is quadratic in x and let*

$$\Phi_p^x \leq \left(1 - \frac{1}{\delta}\right)\frac{1}{2 - \delta}$$

for some $\delta \in (1,2)$. Now let $q(x,y) = (1-\partial_x)p(x+\delta, y)$ and assume that (x_0, y_0) is above the roots of both p and q. Then

$$\Phi_q^y \leq \Phi_p^y.$$

Proof. We first write q as $R(p)$ for a differential operator R. The shift by δ can be translated into a differential operator using Taylor's formula:

$$f(x+t) = \sum_i f^{(i)}(x)\frac{t^i}{i!} = \sum_i \frac{t^i}{i!}\partial_x^i f. \tag{3.1}$$

Hence we can write

$$\begin{aligned} q(x,y) &= p(x+\delta, y) - p_x(x+\delta, y) \\ &= \left(1 + \delta\partial_x + \frac{\delta^2}{2}\partial_x^2\right) p(x,y) - \left(\partial_x + \delta\partial_x^2\right) p(x,y) \\ &= \left(1 + (\delta-1)\partial_x + \left(\frac{\delta^2}{2} - \delta\right)\partial_x^2\right) p \end{aligned}$$

where all higher level derivatives can be discarded since p is quadratic. Hence we can write

$$q(x,y) = a_0 p + a_1 p_x + a_2 p_{xx}$$

where $a_0 = 1$ and

$$a_1 = \delta - 1 \quad \text{and} \quad a_2 = \frac{\delta^2}{2} - \delta. \tag{3.2}$$

Then Lemma 3.6 implies

$$\Phi_q^y \leq \Phi_p^y$$

if and only if

$$a_1 \tilde{D}\left(\hat{L}\right) + 2a_2 \tilde{D}\left(\hat{A}, \hat{L}\right) \geq 0. \tag{3.3}$$

Since $\hat{L}$ is positive semidefinite, $\tilde{D}\left(\hat{L}\right) = 0$ if and only if $\hat{L}$ is the 0 matrix. In this case, we would also have $\tilde{D}\left(\hat{A}, \hat{L}\right) = 0$ and so (3.3) would hold trivially (thus finishing the proof). On the other hand, if $\tilde{D}\left(\hat{L}\right) \neq 0$ then (3.3) holds if and only if

$$a_1 \geq -2a_2 \frac{\tilde{D}\left(\hat{A}, \hat{L}\right)}{\tilde{D}\left(\hat{L}\right)} \geq 0.$$

By Corollary 3.4, $\hat{A}$ and $\hat{L}$ are positive semidefinite, so by Lemma 2.17 we have the inequality

$$\tilde{D}\left(\hat{A}, \hat{L}\right) \leq \tilde{D}\left(\hat{A}\right)\tilde{D}\left(\hat{L}\right) = \tilde{D}\left(\hat{L}\right)\Phi_p^x.$$

since $\tilde{D}(A) = \Phi_p^x$ by Corollary 3.5. Now since $\delta \in (1,2)$, we have $a_1 > 0$ and $a_2 < 0$, and so it suffices to show

$$a_1 \geq -2a_2\Phi_p^x.$$

Plugging back in the values for a_1 and a_2 from (3.2) gives

$$\delta - 1 \geq \delta(2-\delta)\Phi_p^x$$

which is precisely our initial hypothesis. □

In Section 4.1, we will use Lemma 3.7 to understand how the transformation from q to p changes the value of the barrier function Φ. To do so, however, we will need to ensure that the point (x_0, y_0) is above the roots of the resulting q (something that is not true in general). In [28], this was addressed (for an appropriately chosen point (x_0, y_0)) using [28, Lemma 5.10]. Again we will need a strengthened version that takes advantage of the quadratic nature of our polynomials.

Lemma 3.8. *Let s be a quadratic, univariate polynomial with positive first coefficient and real roots $a \leq b$. Then*

$$f(x) = x - \frac{2s(x)}{s'(x)}$$

is a nonincreasing function for any $x > b$.

Proof. We start by writing

$$f'(x) = 1 - 2\frac{s'(x)^2 - s''(x)s(x)}{s'(x)^2} = 2\frac{s''(x)s(x)}{s'(x)^2} - 1.$$

By Taylor's formula, we have

$$s(x+y) = s(x) + ys'(x) + \frac{y^2}{2}s''(x)$$

which is real-rooted (as a polynomial in y) and therefore by the quadratic formula we must have

$$s'(x)^2 \geq 2s(x)s''(x)$$

with equality if and only if s has a double root. Since s has positive first coefficient, both $s(x)$ and $s''(x)$ are nonnegative for $x \geq b$ — we therefore have

$$f'(x) \leq 0$$

with equality if and only if s has a double root. □

4. Proof of Theorem 1.2

The purpose of this section is to prove Theorem 1.5 using the tools from Sections 2 and 3. Theorem 1.2 can then be deduced from Theorem 1.5 by the same argument as [28, Theorem 1.4] (which we briefly review here).

Given random vectors $v_1, \ldots, v_m \in \mathbb{C}^d$, one can define the (random) matrix $V = \sum_i v_i v_i^*$ and its (random) characteristic polynomial

$$p_V(x) = \det\left[x\mathbf{I} - V\right].$$

In the case that the random vectors $\{v_i\}$ are independent, the authors of [28] constructed a so-called *interlacing family* from the polynomials in the support of p_V. As mentioned in Section 1.1, any such construction provides a polynomial p^* with the following properties:

(1) p^* has all real roots,
(2) $\mathbb{P}\left(\text{maxroot}(p_V) < \text{maxroot}(p^*)\right) > 0$.

Since each p_V is the characteristic polynomial of a positive semidefinite matrix, the largest root of p_V is the operator norm of V. Hence a conclusion like the one in Theorem 1.2 could be obtained by finding an appropriate bound on $\operatorname{maxroot}(p^*)$.

In the case of the interlacing family constructed in [28], the associated polynomial p^* is the expected characteristic polynomial

$$p^*(x) = \mathbb{E}\left[p_V(x)\right] = \mathbb{E}\left[\det\left[x\mathbf{I} - \sum_i v_i v_i^*\right]\right]. \tag{4.1}$$

As the first step in the process of bounding the largest root of p^*, Marcus, Spielman, and Srivastava [28, Theorem 4.1] showed that (4.1) could be written in a form they call a *mixed characteristic polynomial*:

$$p^*(x) = \mathbb{E}\left[\det\left[x\mathbf{I} - \sum_i v_i v_i^*\right]\right] = \left(\prod_{i=1}^m 1 - \partial_{y_i}\right) \det\left[\sum_{i=1}^m y_i A_i\right]\Bigg|_{y_1=\cdots=y_m=x}$$

where $A_i = \mathbb{E}\left[v_i v_i^*\right]$ for each i. By translating the restrictions on the v_i in the hypothesis of Theorem 1.2 to restrictions on the associated A_i, one can see that Theorem 1.5 is precisely the bound on the largest root of p^* necessary for Theorem 1.2.

Once Theorem 1.2 is established, the deduction of Corollary 1.3 then follows the same proof as [28, Corollary 1.5]. For the benefit of the reader, we reproduce it here:

Proof of Corollary 1.3. For each $i \in [m]$ and $k = 1, 2$, we define vectors $w_{i,k} \in \mathbb{C}^{2d}$ by

$$w_{i,1} = \sqrt{2/\eta}\begin{bmatrix} u_i \\ 0^d \end{bmatrix},\ w_{i,2} = \sqrt{2/\eta}\begin{bmatrix} 0^d \\ u_i \end{bmatrix}.$$

Let $v_1, \ldots, v_m$ be independent random vectors such that $\mathbb{P}\left(v_i = w_{i,k}\right) = 1/2$, $k = 1, 2$. A simple calculation using (1.3) shows that these vectors satisfy (1.1) with $\epsilon = 2/\eta < 1/2$. By Theorem 1.2, there exists an assignment of each v_i so that

$$\left\|\sum_{i=1}^m v_i v_i^*\right\| = \left\|\sum_{k=1}^2 \sum_{i \in S_k} w_{i,k} w_{i,k}^*\right\| \leq 1 + 2\sqrt{\epsilon}\sqrt{1-\epsilon},$$

where $I_k = \{i : v_i = w_{i,k}\}$. Hence, for $k = 1, 2$, we have

$$\left\|\sum_{i \in S_k} u_i u_i^*\right\| = \frac{\eta}{2}\left\|\sum_{i \in S_k} w_{i,k} w_{i,k}^*\right\| \leq \frac{\eta}{2}\left(1 + 2\sqrt{\epsilon(1-\epsilon)}\right) = \frac{\eta}{2} + \sqrt{2(\eta-2)}.$$

This shows (1.4) with $\theta = \eta/2 - \sqrt{2(\eta-2)} > 0$ (when $\eta > 4$). □

4.1. **Proof of Theorem 1.5.** We first recall the formal definition of the *barrier function* from [28]:

Definition 4.1. For a polynomial $Q \in \mathbb{C}[x_1, \ldots, x_m]$ and a point $\vec{y} \in \mathbb{R}^m$, we define the *barrier function of Q at $\vec{y}$* to be the function

$$\Phi_Q^i(\vec{y}) = \begin{cases} \frac{\partial}{\partial x_i} \ln Q(\vec{y}) & \text{when } \vec{y} \text{ is above the roots of } Q \\ \infty & \text{otherwise} \end{cases}$$

This is an extension of the function Φ introduced in the previous section to allow for different reference points and more variables (the coordinates x, y have been replaced by variables x_i, and only the subscript i is used so as to reduce the clutter).

Let $\epsilon < 1/2$ and set

$$\delta = 1 + \sqrt{\frac{\epsilon}{1-\epsilon}} \quad \text{and} \quad t = (1-2\epsilon)\sqrt{\frac{\epsilon}{1-\epsilon}}.$$

We start with a polynomial Q_0 and reference point w_0 defined as

$$Q_0 = \det\left[\sum_i x_i A_i\right] \quad \text{and} \quad w_0 = t\vec{1}$$

where $\vec{1}$ is the vector with d 1's. It is easy to check that w_0 is above the roots of Q_0 and that $\Phi^j_{Q_0}(w_0) \leq \frac{\epsilon}{t}$ for all j as in the proof of [28, Theorem 5.1].

Given Q_i and w_i, we will construct polynomial Q_{i+1} and reference point w_{i+1} as

$$Q_{i+1} = (1-\partial_{i+1})Q_i \quad \text{and} \quad w_{i+1} = w_i + \delta e_{i+1}.$$

Our goal is to show that w_{i+1} is above the roots of Q_{i+1}. To do so, we will need to understand the effect of applying the $(1-\partial_{i+1})$ operator on the polynomial Q_i, for which we can use the barrier function $\Phi^{i+1}_{Q_i}(w_i)$:

Lemma 4.2. *If w_i is above the roots of Q_i and*

$$\Phi^{i+1}_{Q_i}(w_i) \leq \frac{\epsilon}{t}$$

then w_{i+1} is above the roots of Q_{i+1}.

Proof. Let $s(x)$ be the univariate polynomial that comes from holding all variables of Q_i other than the $(i+1)$st variable constant, i.e., $s(x) = Q_i(t+\delta, \ldots, t+\delta, x, t, \ldots, t)$. By the monotonicity of barrier functions [28, Lemma 5.8], it suffices to show that

$$s(t+\delta) - s'(t+\delta) > 0 \tag{4.2}$$

given that

$$\Phi_s(t) := \frac{s'(t)}{s(t)} \leq \frac{\epsilon}{t}.$$

Note that equation (4.2) is equivalent to $\Phi_s(t+\delta) < 1$. By Lemma 3.8, we have

$$t - \frac{2}{\Phi_s(t)} \geq t + \delta - \frac{2}{\Phi_s(t+\delta)}$$

and so

$$\frac{1}{\Phi_s(t+\delta)} \geq \frac{\delta}{2} + \frac{1}{\Phi_s(t)}.$$

Thus it suffices to show

$$\frac{\delta}{2} + \frac{1}{\Phi_s(t)} > 1$$

which is equivalent to showing

$$\Phi_s(t) < \frac{2}{2-\delta}.$$

Plugging in the hypothesis, it suffices to show

$$\frac{\epsilon}{t} < \frac{2}{2-\delta}$$

which reduces to showing

$$\frac{\sqrt{\epsilon}\sqrt{1-\epsilon}}{1-2\epsilon} < \frac{2}{1-\sqrt{\frac{\epsilon}{1-\epsilon}}}$$

when the given values of δ and t are inserted. It is then easy to check that this holds for any $\epsilon < 1/2$. □

In order to use Lemma 4.2, we will need to bound the value of $\Phi^{i+1}_{Q_i}(w_i)$. We will do this by showing that the transformation from (Q_i, w_i) to (Q_{i+1}, w_{i+1}) causes the barrier functions to shrink in all coordinates $j > i+1$. Note that when moving from (Q_i, w_i) to (Q_{i+1}, w_{i+1}), we are altering only the x_{i+1} variable. Hence to see what happens to the barrier function in coordinate $j > i+1$, we can restrict to those two coordinates (since the restriction of a real stable polynomial is a real stable polynomial) and appeal to Lemma 3.7.

Lemma 4.3. *Let j be a coordinate such that $i+1 < j \leq m$ and*

$$\Phi^j_{Q_i}(w_i) \leq \frac{\epsilon}{t}.$$

Then

$$of\Phi^j_{Q_{i+1}}(w_{i+1}) \leq \frac{\epsilon}{t}.$$

Proof. Lemma 4.2 ensures that w_{i+1} is above the roots of Q_{i+1} and so it is sufficient (by Lemma 3.7) to show

$$\Phi^j_{Q_i}(w_i) \leq \left(1 - \frac{1}{\delta}\right)\frac{1}{2-\delta}.$$

Using the hypothesis, this would be implied by showing

$$\frac{\epsilon}{t} \leq \left(1 - \frac{1}{\delta}\right)\frac{1}{2-\delta}.$$

However one can easily check that

$$\frac{\epsilon}{t} = \frac{\sqrt{\epsilon}\sqrt{1-\epsilon}}{1-2\epsilon} = \left(1 - \frac{1}{\delta}\right)\frac{1}{2-\delta}$$

and so we are done. □

Iterating m times, Lemmas 4.2 and 4.3 ensure that w_m is above the roots of Q_m, where Q_m is exactly the polynomial in Equation (1.5) before the variables are set to x. Furthermore, $w_m = (t+\delta)\vec{1}$. Hence w_m being above the roots of Q_m implies the largest root of Equation (1.5) is at most

$$t + \delta = 1 + 2\sqrt{\epsilon}\sqrt{1-\epsilon}$$

as required for Theorem 1.5.

Remark 4.4. The argument here is more delicate than the one given in [28]; this can be seen by comparing the statement of Lemma 3.7 to its analogous version [28, Lemma 5.11]. In [28], any δ that caused the barrier function to contract also resulted in w_{i+1} being above the roots of Q_{i+1} (a fortiori). This is not the case here and is the reason that the additional hypothesis of (x_0, y_0) being above the roots is necessary in Lemma 3.7. This also becomes evident when considering the space of values (δ, t) for which Lemma 4.2 and Lemma 4.3 hold. In [28], the constraint provided by [28, Lemma 5.11] was the only relevant one in determining the optimal values of δ and t, whereas in our case both Lemma 3.7 and Lemma 3.8 provide nontrivial constraints.

5. Naimark's complements of frame partitions

In this section we establish a result that links properties of two complementary subsets of a Parseval frame with the corresponding subsets of the Naimark's complement. In general, a subset of a Parseval frame does not need to be a frame, and we can only expect it to be a Bessel sequence with bound 1. In Proposition 5.4 we show that if this subset has a Bessel bound strictly less than 1, then its corresponding Naimark's complement subset is a Riesz sequence.

It is rather surprising that Proposition 5.4 has not appeared in the frame theory literature before despite its simplicity and the elementary nature of its proof. However, it can be considered as a quantitative variant of the complementarity principle between spanning and linear independence due to Bodmann, Casazza, Paulsen, and Speegle [12, Proposition 2.3]. We start with basic conventions in frame theory [20].

Definition 5.1. A family of vectors $\{\phi_i\}_{i\in I}$ in a Hilbert space $\mathcal{H}$ is called a *frame* for $\mathcal{H}$ if there are constants $0 < A \leq B < \infty$ (called *lower and upper frame bounds*, respectively) so that

$$A\|\phi\|^2 \leq \sum_{i\in I} |\langle \phi, \phi_i\rangle|^2 \leq B\|\phi\|^2 \qquad \text{for all } \phi \in \mathcal{H}. \tag{5.1}$$

If we only have the right hand inequality in (5.1), we call $\{\phi_i\}_{i\in I}$ a *Bessel sequence* with Bessel bound B. If $A = B$, $\{\phi_i\}_{i\in I}$ is called a *tight frame* and if $A = B = 1$, it is called a *Parseval frame*.

Definition 5.2. A family of vectors $\{\phi_i\}_{i\in I}$ in a Hilbert space $\mathcal{H}$ is a *Riesz sequence* if there are constants $A, B > 0$ so that for all $\{a_i\} \in \ell^2(I)$ we have

$$A\sum_{i\in I} |a_i|^2 \leq \left\| \sum_{i\in I} a_i\phi_i \right\|^2 \leq B\sum_{i\in I} |a_i|^2. \tag{5.2}$$

We call A, B *lower and upper Riesz bounds* for $\{\phi_i\}_{i\in I}$.

Note that it suffices to verify (5.2) only for sequences $\{a_i\}$ with finitely many non-zero coefficients, since a standard convergence argument yields the same bounds (5.2) for all infinitely supported sequences $\{a_i\} \in \ell^2(I)$. In general we do not require that frame, Bessel, and Riesz bounds in Definitions 5.1 and 5.2 are optimal. In particular, a Bessel sequence with bound B is automatically a Bessel sequence with bound $B' \geq B$.

Notation 5.3. Throughout the rest of the paper $\{e_i\}_{i\in I}$ will denote an orthonormal basis for whatever space we are working in.

Proposition 5.4. *Let $P : \ell^2(I) \to \ell^2(I)$ be the orthogonal projection onto a closed subspace $\mathcal{H} \subset \ell^2(I)$. Then, for any subset $J \subset I$ and $\delta > 0$, the following are equivalent:*

(i) $\{Pe_i\}_{i\in J}$ *is a Bessel sequence with bound* $1-\delta$,
(ii) $\{Pe_i\}_{i\in J^c}$ *is a frame with lower bound* δ,
(iii) $\{(\mathbf{I}-P)e_i\}_{i\in J}$ *is a Riesz sequence with lower bound* δ, *where* $\mathbf{I}$ *is the identity on* $\ell^2(I)$.

Proof. Since $\{Pe_i\}_{i\in I}$ is a Parseval frame for $\mathcal{H}$, we have

$$\|\phi\|^2 = \sum_{i\in J} |\langle \phi, Pe_i\rangle|^2 + \sum_{i\in J^c} |\langle \phi, Pe_i\rangle|^2 \qquad \text{for all } \phi \in \mathcal{H}. \tag{5.3}$$

Thus,

$$\sum_{i\in J} |\langle \phi, Pe_i\rangle|^2 \le (1-\delta)\|\phi\|^2 \iff \sum_{i\in J^c} |\langle \phi, Pe_i\rangle|^2 \ge \delta\|\phi\|^2 \qquad \text{for all } \phi \in \mathcal{H}. \tag{5.4}$$

By (5.3), the Bessel sequence $\{Pe_i\}_{i\in J^c}$ inherits the Bessel bound 1 as a subset of a Parseval frame. Thus, (5.4) shows the equivalence of (i) and (ii).

To show the equivalence of (i) and (iii), note that for any sequence of coefficients $\{a_i\} \in \ell^2(J)$,

$$\sum_{i\in J} |a_i|^2 = \left\| \sum_{i\in J} a_i Pe_i \right\|^2 + \left\| \sum_{i\in J} a_i (\mathbf{I}-P)e_i \right\|^2. \tag{5.5}$$

Thus,

$$\left\| \sum_{i\in J} a_i Pe_i \right\|^2 \le (1-\delta) \sum_{i\in J} |a_i|^2 \iff \left\| \sum_{i\in J} a_i (\mathbf{I}-P)e_i \right\|^2 \ge \delta \sum_{i\in J} |a_i|^2. \tag{5.6}$$

By (5.5), the family $\{(\mathbf{I}-P)e_i\}_{i\in J}$ has automatically the Riesz upper bound 1. Observe that the inequality in left hand side of (5.6) is equivalent to (i). This follows from the well-known fact that adjoint of the analysis operator

$$T : \mathcal{H} \to \ell^2(I), \qquad T\phi = \{\langle \phi, Pe_i\rangle\}_{i\in J}, \quad \phi \in \mathcal{H},$$

is the synthesis operator

$$T^* : \ell^2(I) \to \mathcal{H}, \qquad T^*(\{a_i\}_{i\in J}) = \sum_{i\in J} a_i Pe_i, \quad \{a_i\}_{i\in J} \in \ell^2(J).$$

Since $||T|| = ||T^*||$, (5.6) yields the equivalence of (i) and (iii). □

Remark 5.5. A curious reader might ask what condition needs to be imposed about $\{(\mathbf{I}-P)e_i\}_{i\in J^c}$ to obtain the equivalence in Proposition 5.4. Surprisingly, this condition can not be easily stated in terms of Bessel, Riesz, or frame bounds. Instead, it is not difficult to show that the following restricted Riesz upper bound condition does the job:

$$\left\| \sum_{i\in J^c} \langle \phi, Pe_i\rangle (\mathbf{I}-P)e_i \right\|^2 \le (1-\delta) \sum_{i\in J^c} |\langle \phi, Pe_i\rangle|^2 \qquad \text{for all } \phi \in \mathcal{H}. \tag{iv}$$

Since this observation will not be used subsequently in the paper, we leave the verification of details to the reader.

As an immediate consequence of Proposition 5.4 we have

Corollary 5.6. *Let $P : \ell^2(I) \to \ell^2(I)$ be the orthogonal projection onto a closed subspace $\mathcal{H} \subset \ell^2(I)$. Then, for any subset $J \subset I$ and $\delta > 0$, the following are equivalent:*

(i) $\{Pe_i\}_{i\in J}$ *is a frame with frame bounds* δ *and* $1-\delta$,
(ii) $\{Pe_i\}_{i\in J^c}$ *is a frame with frame bounds* δ *and* $1-\delta$,
(iii) both $\{Pe_i\}_{i\in J}$ *and* $\{Pe_i\}_{i\in J^c}$ *are Bessel sequences with bound* $1-\delta$,
(iv) both $\{(\mathbf{I}-P)e_i\}_{i\in J}$ *and* $\{(\mathbf{I}-P)e_i\}_{i\in J^c}$ *are Riesz sequences with lower bound* δ.

Proof. Suppose that (iii) holds. Applying Proposition 5.4 simultaneously to Bessel sequences $\{Pe_i\}_{i\in J}$ and $\{Pe_i\}_{i\in J^c}$ yields the remaining properties (i), (ii), and (iv). Conversely, any of these properties implies (iii) in light of Proposition 5.4. □

6. Asymptotic bounds in Feichtinger Conjecture

In this section we establish quantitative bounds in Feichtinger Conjecture. To achieve this we shall employ the results of the previous section and the landmark result of Marcus, Spielman, and Srivastava [28, Corollary 1.5]. In the language of Bessel sequences this result takes the following form, where $\mathcal{H}_n$ denotes n-dimensional real or complex Hilbert space $\mathbb{R}^n$ or $\mathbb{C}^n$.

Theorem 6.1. *Let $\{u_i\}_{i=1}^M \subset \mathcal{H}_n$ be a Bessel sequence with bound 1 and $\|u_i\|^2 \leq \delta$ for all i. Then for any positive integer r, there exists a partition $\{I_1, \dots, I_r\}$ of $[M]$ such that each $\{u_i\}_{i\in I_j}$, $j = 1, \dots, r$, is a Bessel sequence with bound*

$$\left(\frac{1}{\sqrt{r}} + \sqrt{\delta}\right)^2.$$

Remark 6.2. Note that the original formulation [28] of Theorem 6.1 requires that $\{u_i\}_{i=1}^M$ is Parseval frame. This can be relaxed since any Bessel sequence $\{u_i\}_{i=1}^M$ with bound 1 can be extended to a Parseval frame by adding additional vectors satisfying $\|u_i\|^2 \leq \delta$ for $i = M+1, \dots, \tilde{M}$, $\tilde{M} > M$. This is a consequence of the Schur–Horn Theorem [5, 13], see also the proof of Corollary 6.6.

Corollary 1.3 gives us a quantitative version of Weaver Conjecture KS_2 with sharper constants than those deducible from Theorem 6.1. In particular, a simple rescaling of Corollary 1.3 combined with Remark 6.2 yields the following theorem, which we state for Bessel sequences.

Theorem 6.3. *Let $0 < \delta_0 < 1/4$ and $\varepsilon_0 = 1/2 - \sqrt{2\delta_0(1-2\delta_0)}$. Suppose that $\{\phi_i\}_{i=1}^M$ is a Bessel sequence in $\mathcal{H}_n$ with Bessel bound 1 and that $\|\phi_i\|^2 \leq \delta_0$ for all i. Then there exists a partition $\{I_1, I_2\}$ of $[M]$ such that each $\{\phi_i\}_{i\in I_j}$, $j = 1, 2$, is a Bessel sequence with bound $1-\varepsilon_0$.*

Notation 6.4. Since the above constants might potentially be improved in the future, we shall keep δ_0 and ϵ_0 as base parameters that shall propagate to all remaining results in this section. Consequently, we shall fix the constants δ_0 and ε_0 as in Theorem 6.3 throughout this section. In particular, combining Corollary 5.6 with Theorem 6.3 yields the following result with the same constants δ_0 and ϵ_0.

Corollary 6.5. *Suppose that $\{\phi_i\}_{i=1}^M$ is a Parseval frame for $\mathcal{H}_n$ and $\|\phi_i\|^2 \geq 1-\delta_0$ for all i. Then there exists a partition $\{I_1, I_2\}$ of $[M]$ such that each $\{\phi_i\}_{i\in I_j}$, $j = 1, 2$, is a Riesz sequence with lower bound ε_0.*

Proof. By Naimark's Theorem we can write $\phi_i = Pe_i$, where P is an orthogonal projection of $\mathcal{H}_M$ onto n-dimensional subspace identified with $\mathcal{H}_n$. Applying Theorem 6.3 to $\{(\mathbf{I}-P)e_i\}_{i=1}^M$ yields a partition $\{I_1, I_2\}$ of $[M]$ such that each $\{(\mathbf{I}-P)e_i\}_{i\in I_j}$, $j = 1, 2$, is a Bessel sequence with bound $1-\varepsilon_0$. By Corollary 5.6, each $\{Pe_i\}_{i\in I_j}$, $j = 1, 2$, is a Riesz sequence with lower bound ε_0. $\square$

Next, we extend the validity of Corollary 6.5 to Bessel sequences. This requires a more sophisticated approach than what was outlined in Remark 6.2.

Corollary 6.6. *Suppose that $\{\phi_i\}_{i=1}^M$ is a Bessel sequence with bound B for $\mathcal{H}_n$ and $\|\phi_i\|^2 \geq B(1-\delta_0)$ for all i. Then there exists a partition $\{I_1, I_2\}$ of $[M]$ such that each $\{\phi_i\}_{i\in I_j}$, $j = 1, 2$, is a Riesz sequence with lower bound $B\varepsilon_0$.*

Proof. Without loss of generality we can assume that the Bessel bound $B = 1$. Suppose that the frame operator of $\{\phi_i\}_{i=1}^M$, which is given by

$$S = \sum_{i=1}^{M} \phi_i \phi_i^*,$$

has eigenvalues $1 \geq \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$. For a fixed N, consider an operator $\tilde{S} = \mathbf{I}_{n+N} - S \oplus \mathbf{0}_N$ on $\mathcal{H}_{n+N} = \mathcal{H}_n \oplus \mathcal{H}_N$, where $\mathbf{0}_N$ is the zero operator on $\mathcal{H}_N$. Then, $\tilde{S}$ has the following eigenvalues listed in decreasing order

$$\underbrace{1, \ldots, 1}_{N}, 1-\lambda_n, \ldots, 1-\lambda_1. \tag{6.1}$$

Our goal is to find a collection of vectors $\{\phi_{M+i}\}_{i=1}^K$ in $\mathcal{H}_{n+N}$ such that:

(i) its frame operator is $\tilde{S}$, and
(ii) $\|\phi_{M+i}\|^2 = C$ for all $i = 1, \ldots, K$ for some constant $C \in [1-\delta_0, 1]$.

By the Schur–Horn Theorem [5, 13] this is possible if and only if the sequence (6.1) majorizes

$$\underbrace{C, \ldots, C}_{K}.$$

This, in turn, is implied by

$$K \geq N + n \qquad \text{and} \qquad CK = N + \sum_{i=1}^{n} (1-\lambda_i). \tag{6.2}$$

By choosing sufficiently large N and $K = N + n$, we have that

$$1 - \delta_0 \leq C = \frac{N + \sum_{i=1}^{n}(1-\lambda_i)}{N+n} \leq 1.$$

This shows the existence of vectors $\{\phi_{M+i}\}_{i=1}^K$ satisfying (i) and (ii).

Now, $\{\phi_i\}_{i=1}^{M+K}$ is a Parseval frame for $\mathcal{H}_{n+N}$ such that $\|\phi_i\|^2 \geq 1-\delta_0$ for all i. By Corollary 6.5 there exists a partition $\tilde{I}_1$, $\tilde{I}_2$ of $[M+K]$ such that each $\{\phi_i\}_{i\in\tilde{I}_j}$, $j = 1, 2$, is a Riesz sequence with lower bound ε_0. Then, $I_j = \tilde{I}_j \cap [M]$ is the required partition of $[M]$. $\square$

We are now ready to show the asymptotic estimate on the size of a partition in the Feichtinger Conjecture.

Theorem 6.7. *Suppose that $\{\phi_i\}_{i=1}^M$ is a Bessel sequence with bound 1 for $\mathcal{H}_n$ and $\|\phi_i\|^2 \geq \varepsilon > 0$ for all i. Then there exists $r = r(\varepsilon) = O(1/\varepsilon)$ and a partition $\{I_1, \ldots, I_r\}$ of $[M]$ such that each $\{\phi_i\}_{i\in I_j}$, $j = 1, \ldots, r$, is a Riesz sequence with lower bound $\varepsilon_0\varepsilon$. In addition, if $\|\phi_i\|^2 = \varepsilon$ for all i, then the upper Riesz bound of each $\{\phi_i\}_{i\in I_j}$ is $\frac{\varepsilon}{1-\delta_0}$.*

Proof. First, observe that without loss of generality we can assume that $\|\phi_i\|^2 = \varepsilon$ for all i. Indeed, we can replace each ϕ_i by $\frac{\sqrt{\varepsilon}}{\|\phi_i\|}\phi_i$ and then apply Theorem 6.7 to get the general case.

By Theorem 6.1 for each $\tilde{r}$ we can find a partition $\{\tilde{I}_j\}_{j=1}^{\tilde{r}}$ of $[M]$ such that each $\{\phi_i\}_{i\in\tilde{I}_j}$ is a Bessel sequence with bound

$$B = \left(\frac{1}{\sqrt{\tilde{r}}} + \sqrt{\varepsilon}\right)^2.$$

Then we wish to apply Corollary 6.6 to each such $\{\phi_i\}_{i\in\tilde{I}_j}$. This is possible if we choose $\tilde{r}$ such that

$$\|\phi_i\|^2 = \varepsilon \geq B(1-\delta_0) = \left(\frac{1}{\sqrt{\tilde{r}}} + \sqrt{\varepsilon}\right)^2 (1-\delta_0). \tag{6.3}$$

A simple calculation shows that the above inequality simplifies to

$$\frac{2}{\sqrt{\tilde{r}\varepsilon}} + \frac{1}{\tilde{r}\varepsilon} \leq \frac{\delta_0}{1-\delta_0}.$$

Hence, it suffices to choose

$$\tilde{r} \geq \frac{9}{\varepsilon}\left(\frac{1-\delta_0}{\delta_0}\right)^2.$$

By Corollary 6.6, each $\{\phi_i\}_{i\in\tilde{I}_j}$ can be partitioned into two 2 Riesz sequences with lower bound $B\varepsilon_0 \geq \varepsilon_0\varepsilon$ and upper bound $B \leq \frac{\varepsilon}{1-\delta_0}$. This gives the required partition of size $r = 2\tilde{r}$ and completes the proof of Theorem 6.7. □

Remark 6.8. Note that choosing $\tilde{r}$ so that (6.3) is almost an equality yields the slightly better lower bound $\frac{\varepsilon_0}{1-\delta_0}\varepsilon$ in the conclusion of Theorem 6.7. Moreover, since we can take any $\delta_0 < 1/4$, we obtain an explicit estimate on the size of a partition $r = r(\varepsilon) \leq 162/\varepsilon$. Surely, the number 162 is far from sharp, but it merely gives a crude upper bound on optimal constant.

6.1. **Infinite dimensional results.** Corollary 1.3 can be easily extended to the infinite dimensional setting using the "pinball principle" [15, Proposition 2.1], which we state here for the reader's convenience.

Theorem 6.9. *Fix a natural number r and assume for every natural number n, we have a partition $\{I_i^n\}_{i=1}^r$ of $[n]$. Then there are natural numbers $\{n_1 < n_2 < \cdots\}$ so that if $j \in I_i^{n_j}$ for some $i \in [r]$, then $j \in I_i^{n_k}$ for all $k \geq j$. For any $i \in [r]$ define $I_i = \{j : j \in I_i^{n_j}\}$. Then,*

(i) $\{I_i\}_{i=1}^r$ is a partition of $\mathbb{N}$.

(ii) If $I_i = \{j_1 < j_2 < \cdots\}$, then for every natural number k we have $\{j_1, j_2, \ldots, j_k\} \subset I_i^{n_{j_k}}$.

Theorem 6.10. *Suppose $\{\phi_i\}_{i\in I}$ is a Bessel sequence in a separable Hilbert space $\mathcal{H}$ with constant $\eta > 4$, which consists of vectors of norms $\|\phi_i\| \le 1$. That is,*

$$\sum_{i\in I} |\langle \phi, \phi_i\rangle|^2 \le \eta \|\phi\|^2 \qquad \text{for all } \phi \in \mathcal{H}.$$

Then there exists a constant $\theta = \theta(\eta) > 0$, depending only on $\eta > 4$, such that the index set I can be decomposed into subsets I_1 and I_2, so that the Bessel sequences $\{\phi_i\}_{i\in I_j}$, $j = 1, 2$, have bounds $\le \eta - \theta$, i.e.,

$$\sum_{i\in I_j} |\langle \phi, \phi_i\rangle|^2 \le (\eta - \theta)\|\phi\|^2 \qquad \text{for all } \phi \in \mathcal{H},\ j = 1, 2.$$

Proof. Without loss of generality, we can assume that the index set $I = \mathbb{N}$. For each $n \in \mathbb{N}$, the set $\{\phi_i\}_{i=1}^n$ is a Bessel sequence in $\mathcal{H}$ with Bessel bound η. By Theorem 6.3, for each n, there exists a partition $\{I_1^n, I_2^n\}$ of $[n]$ such that for each $j = 1, 2$, $\{\phi_i\}_{i\in I_j^n}$ is a Bessel sequence with bound $\eta - \theta$. Let $\{I_1, I_2\}$ be the partition of $\mathbb{N}$ obtained by applying Theorem 6.9 to the sequence of partitions $\{I_1^n, I_2^n\}_{n=1}^\infty$. For any $j = 1, 2$, we write $I_j = \{j_1 < j_2 < \cdots\}$. By Theorem 6.9(ii) for any $k \in \mathbb{N}$, $\{\phi_{j_1}, \ldots, \phi_{j_k}\}$ is a Bessel sequence with bound $\eta - \theta$. Since $k \in \mathbb{N}$ is arbitrary, $\{\phi_i : i \in I_j\}$ is a Bessel sequence with bound $\eta - \theta$ for $j = 1, 2$, as desired. □

In a similar way, the "pinball principle" implies the infinite dimensional Theorem 6.11 from finite dimensional Theorem 6.7. Theorem 6.11 is simply a rescaled variant of Theorem 1.4.

Theorem 6.11. *Suppose that $\{\phi_i\}_{i\in I}$ is a Bessel sequence with bound B for a separable Hilbert space $\mathcal{H}$ and $\|\phi_i\| \ge 1$ for all i. Then there exists $r = O(B)$ and a partition $\{I_1, \ldots, I_r\}$ of I such that each $\{\phi_i\}_{i\in I_j}$, $j = 1, \ldots, r$, is a Riesz sequence with lower bound ε_0. Moreover, if $\|\phi_i\| = 1$ for all i, then the upper bound is $\frac{1}{1-\delta_0}$.*

Proof. Without loss of generality we can assume that the index set $I = \mathbb{N}$ and $\|\phi_i\| = 1$ for all i. For every $n \in \mathbb{N}$, we apply Theorem 6.7 to the Bessel sequence $\{\frac{1}{\sqrt{B}}\phi_i\}_{i=1}^n$ to obtain a partition $\{I_1^n, \ldots, I_r^n\}$ of $[n]$ such that each subsequence $\{\frac{1}{\sqrt{B}}\phi_i\}_{i\in I_j^n}$ is a Riesz sequence with bounds ε_0/B and $\frac{1}{B(1-\delta_0)}$. The size of this partition is some fixed $r = O(B)$ as $B \to \infty$. Let $\{I_1, \ldots, I_r\}$ be the partition of $\mathbb{N}$ obtained by applying Theorem 6.9 to the sequence of partitions $\{I_1^n, \ldots, I_r^n\}_{n=1}^\infty$. For any $j \in [r]$, we write $I_j = \{j_1 < j_2 < \cdots\}$. By Theorem 6.9(ii) for any $k \in \mathbb{N}$, $\{\phi_{j_1}, \ldots, \phi_{j_k}\}$ is a Riesz sequence with bounds ε_0 and $\frac{1}{1-\delta_0}$. Since $k \in \mathbb{N}$ is arbitrary, we conclude that $\{\phi_i\}_{i\in I_j}$ is a Riesz sequence with the same bounds for every $j = 1, \ldots, r$. □

It is also worth investigating the size of a partition if we insist on having nearly tight Riesz sequences. That is, Riesz sequences with bounds $1-\varepsilon$ and $1+\varepsilon$, where $\varepsilon > 0$ is an arbitrarily small parameter. That is, we are asking for the size of partition in R_ε Conjecture that was shown by Casazza and Tremain [19] to be equivalent with the Kadison–Singer problem.

Theorem 6.12. *Suppose that $\{\phi_i\}_{i\in I}$ is a unit norm Bessel sequence with bound B for a separable Hilbert space $\mathcal{H}$. Then for any $\varepsilon > 0$ there exists a partition $\{I_1, \ldots, I_r\}$ of I of size $r = O(B/\varepsilon^4)$, such that each $\{\phi_i\}_{i\in I_j}$, $j = 1, \ldots, r$, is a Riesz sequence with bounds $1-\varepsilon$ and $1+\varepsilon$.*

In the proof of Theorem 6.12 we will use the following lemma. The case when $J = I$ is a well-known fact, see [20, Section 3.6]. For the sake of completeness we will give the proof of Lemma 6.13.

Lemma 6.13. *Suppose $\{\phi_i\}_{i\in I}$ is a Riesz basis in a Hilbert space $\mathcal{H}$, and let $\{\phi_i^*\}_{i\in I}$ be its unique biorthogonal (dual) Riesz basis, i.e.,*

$$\langle \phi_i, \phi_j^* \rangle = \delta_{i,j} \qquad \text{for all } i, j \in I.$$

Then, for any subset $J \subset I$, the Riesz sequence bounds of $\{\phi_i\}_{i\in J}$ are A and B if and only if the Riesz sequence bounds of $\{\phi_i^*\}_{i\in J}$ are $1/B$ and $1/A$.

Proof. Suppose that $\{\phi_i\}_{i\in J}$ has upper Riesz bound B. This is equivalent to the Bessel condition

$$\sum_{i\in J} |\langle f, \phi_i \rangle|^2 \le B||f||^2 \qquad \text{for all } f \in \mathcal{H}. \tag{6.4}$$

For any sequence $\{a_i\}_{i\in J} \in \ell^2$, there exists a unique $f \in \mathcal{H}$ such that

$$\langle f, \phi_i \rangle = \begin{cases} a_i & i \in J, \\ 0 & \text{otherwise.} \end{cases}$$

Since $f = \sum_{i\in J} a_i \phi_i^*$, by (6.4) we have

$$\left\| \sum_{i\in J} a_i \phi_i^* \right\|^2 = ||f||^2 \ge \frac{1}{B} \sum_{i\in J} |\langle f, \phi_i \rangle|^2 = \frac{1}{B} \sum_{i\in J} |a_i|^2.$$

Conversely, if $\{\phi_i^*\}_{i\in J}$ has lower Riesz bound $1/B$, then (6.4) holds and $\{\phi_i\}_{i\in J}$ has upper Riesz bound B. By symmetry, $\{\phi_i^*\}_{i\in J}$ has upper Riesz bound $1/A$ if and only if $\{\phi_i\}_{i\in J}$ has lower Riesz bound A, which completes the proof of the lemma. □

Proof of Theorem 6.12. By the "pinball principle" as in the proof of Theorems 6.10 and 6.11 it suffices to restrict our attention to the finite dimensional case.

In the first step we apply Theorem 6.11 to find a partition of size $O(B)$ into Riesz sequences with bounds ε_0 and $\frac{1}{1-\delta_0}$. Suppose that $\{\phi_i\}_{i=1}^M$ is one of these, i.e., a unit-norm Riesz sequence with bounds ε_0 and $\frac{1}{1-\delta_0}$. Let $\{\phi_i^*\}_{i=1}^M$ be its unique biorthogonal (dual) Riesz sequence in $\mathcal{H}' = \operatorname{span}\{\phi_i\}_{i=1}^M$ that has bounds $1-\delta_0$ and $\frac{1}{\varepsilon_0}$ by Lemma 6.13. In the second step we apply Theorem 6.1 to both of these Riesz sequences to reduce their upper Riesz bounds to $1+\varepsilon$. This requires partitions of size $\tilde{r} = O(1/\varepsilon^2)$ since we need to guarantee that

$$\left(\frac{1}{\sqrt{\tilde{r}}} + 1 \right)^2 \le 1 + \varepsilon.$$

Now it suffices to consider a common refinement partition $I_1, \dots, I_r$ of size $r = O(B/\varepsilon^4)$ of a partition in the first step and two partitions in the second step. For any $j = 1, \dots, r$, $\{\phi_i\}_{i\in I_j}$ is a Riesz sequence with upper bound $1+\varepsilon$. Since $\{\phi_i^*\}_{i\in I_j}$ is also a Riesz sequence with the same upper bound, Lemma 6.13 implies that $\{\phi_i\}_{i\in I_j}$ has lower bound $1/(1+\varepsilon) \ge 1-\varepsilon$. This completes the proof of Theorem 6.12. □

Remark 6.14. It is an open problem what is the optimal dependence of the size of the partition on ε. The linear dependence on B is optimal, but it is less clear whether one can reduce dependence on ε from $O(1/\varepsilon^4)$ to some lower exponent. This problem is closely related with finding the optimal size of partition in Anderson's Paving Conjecture, see [28, Theorem 6.1]. It is known [17, Theorem 6] that size of partition must be at least $O(1/\varepsilon^2)$ as $\varepsilon \to 0$.

Repeating the standard arguments as in [15], Theorem 6.11 yields the same asymptotic bounds on the size of partition for Bourgain–Tzafriri Conjecture.

Theorem 6.15. *There exist universal constants $c, \varepsilon_0, \delta_0 > 0$ so that for any $B > 1$ the following holds. Suppose $T : \mathcal{H} \to \mathcal{H}$ is a linear operator with norm $\|T\| \leq \sqrt{B}$ and $\|Te_i\| = 1$ for all $i \in I$, where $\{e_i\}_{i\in I}$ is an orthonormal basis of a separable Hilbert space $\mathcal{H}$. Then, there exists a partition $\{I_1, \ldots, I_r\}$ of the index set I of size $r \leq cB$, so that for all $j = 1, \ldots, r$ and all choice of scalars $\{a_i\}_{i\in I_j} \in \ell^2$ we have*

$$\varepsilon_0 \sum_{i\in I_j} |a_i|^2 \leq \left\| \sum_{i\in I_j} a_i T e_i \right\|^2 \leq \frac{1}{1-\delta_0} \sum_{i\in I_j} |a_i|^2.$$

As a consequence of our results we obtain explicit bounds on the partition size for Fourier frames. If $E \subset [0,1]$ has positive Lebesgue measure, then the collection of functions $\phi_n(t) = e^{2\pi i n t}\chi_E(t)$, $n \in \mathbb{Z}$, is a Parseval frame for $L^2(E)$, often called a Fourier frame. Since this is an equal norm frame, i.e., $||\phi_n||^2 = |E|$ for all $n \in \mathbb{Z}$, Theorem 6.11 yields the following corollary. Moreover, by the results of Lawton [25] and Paulsen [29, Theorem 1.2], the index sets can be chosen to be a syndetic set. Recall that $I \subset \mathbb{Z}$ is syndetic if for some finite set $F \subset \mathbb{Z}$ we have

$$\bigcup_{n\in F} (I - n) = \mathbb{Z}.$$

Corollary 6.16. *There exists a universal constant $c > 0$ such that for any subset $E \subset [0,1]$ with positive measure, the corresponding Fourier frame $\{\phi_n\}_{n\in\mathbb{Z}}$ can be decomposed as the union of $r \leq c/|E|$ Riesz sequences $\{\phi_n\}_{n\in I_j}$, $j = 1, \ldots, r$. Moreover, each index set I_j can be chosen to be a syndetic set.*

References

1. C. Akemann, J. Anderson, *Lyapunov theorems for operator algebras*, Mem. Amer. Math. Soc. **94** (1991), no. 458, iv+88 pp.
2. J. Anderson, *Extreme points in sets of positive linear maps on $\mathcal{B}(\mathcal{H})$*, J. Funct. Anal. **31** (1979), no. 2, 195–217.
3. J. Anderson, *Extensions, restrictions, and representations of states on C^*-algebras*, Trans. Amer. Math. Soc. **249** (1979), no. 2, 303–329.
4. J. Anderson, *A conjecture concerning the pure states of $\mathcal{B}(\mathcal{H})$ and a related theorem.* Topics in modern operator theory (Timişoara/Herculane, 1980), pp. 27–43, Operator Theory: Adv. Appl., 2, Birkhäuser, Basel-Boston, Mass., 1981.
5. J. Antezana, P. Massey, M. Ruiz, D. Stojanoff, *The Schur-Horn theorem for operators and frames with prescribed norms and frame operator*, Illinois J. Math. **51** (2007), 537–560.
6. S. Artstein-Avidan, D. Florentin and Y. Ostrover, *Remarks about mixed discriminants and volumes*, Commun. Contemp. Math. **16** (2014), no. 2, 1350031, 14 pp.
7. R. B. Bapat, *Mixed discriminants of positive semidefinite matrices*, Linear Algebra Appl. **126** (1989), 107–124.

8. J. Borcea, P. Brändén, *Multivariate Pólya-Schur classification problems in the Weyl algebra*, Proc. Lond. Math. Soc. **101** (2010), 73–104.
9. J. Bourgain, L. Tzafriri, *Invertibility of "large" submatrices with applications to the geometry of Banach spaces and harmonic analysis*, Israel J. Math. **57** (1987), no. 2, 137–224.
10. J. Bourgain, L. Tzafriri, *Restricted invertibility of matrices and applications*, Analysis at Urbana, Vol. II (Urbana, IL, 1986–1987), 61–107, London Math. Soc. Lecture Note Ser., 138, Cambridge Univ. Press, Cambridge, 1989.
11. J. Bourgain, L. Tzafriri, *On a problem of Kadison and Singer*, J. Reine Angew. Math. **420** (1991), 1–43.
12. B. Bodmann, P. Casazza, V. Paulsen, D. Speegle, *Spanning and independence properties of frame partitions*, Proc. Amer. Math. Soc. **140** (2012), no. 7, 2193–2207.
13. M. Bownik, J. Jasper, *Existence of frames with prescribed norms and frame operator*, Excursions in harmonic analysis. Volume 4, 103–117, Appl. Numer. Harmon. Anal., Birkhäuser/Springer, New York, 2015.
14. M. Bownik, D. Speegle, *The Feichtinger conjecture for wavelet frames, Gabor frames and frames of translates*, Canad. J. Math. **58** (2006), no. 6, 1121–1143.
15. P. Casazza, O. Christensen, A. Lindner, R. Vershynin, *Frames and the Feichtinger conjecture*, Proc. Amer. Math. Soc. **133** (2005), no. 4, 1025–1033.
16. P. Casazza, D. Edidin, *Equivalents of the Kadison–Singer problem*, Function spaces, 123–142, Contemp. Math., 435, Amer. Math. Soc., Providence, RI, 2007.
17. P. Casazza, D. Edidin, D. Kalra, V. Paulsen, *Projections and the Kadison–Singer problem*, Oper. Matrices **1** (2007), no. 3, 391–408.
18. P. Casazza, M. Fickus, J. Tremain, E. Weber, *The Kadison–Singer problem in mathematics and engineering: a detailed account*, Operator theory, operator algebras, and applications, 299–355, Contemp. Math., 414, Amer. Math. Soc., Providence, RI, 2006.
19. P. Casazza, J. Tremain, *The Kadison–Singer problem in mathematics and engineering*, Proc. Natl. Acad. Sci. USA **103** (2006), no. 7, 2032–2039.
20. O. Christensen, *An introduction to frames and Riesz bases*, Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc., Boston, MA, 2003.
21. D. Florentin, V. D. Milman, R. Schneider, *A characterization of the mixed discriminant*, Proc. Amer. Math. Soc. **144** (2016), no. 5, 2197–2204.
22. K. Gröchenig, *Localized frames are finite unions of Riesz sequences*, Adv. Comput. Math. **18** (2003), no. 2–4, 149–157.
23. L. Gurvits, *The van der Waerden conjecture for mixed discriminants*, Adv. in Math. **200** (2006), no. 2, 435–454.
24. J. W. Helton, V. Vinnikov, *Linear matrix inequality representation of sets*, Comm. Pure Appl. Math. **60** (2007), no. 5, 654–674.
25. W. Lawton, *Minimal sequences and the Kadison–Singer problem*, Bull. Malays. Math. Sci. Soc. **33** (2010), 169–176.
26. R. Kadison, I. Singer, *Extensions of pure states*, Amer. J. Math. **81** (1959), 383–400.
27. A. W. Marcus, D. A. Spielman, N. Srivastava, *Interlacing families I: bipartite Ramanujan graphs of all degrees*, Ann. of Math. **182** (2015), no. 1, 307–325.
28. A. W. Marcus, D. A. Spielman, N. Srivastava, *Interlacing Families II: mixed characteristic polynomials and the Kadison–Singer problem*, Ann. of Math. **182** (2015), no. 1, 327–350.
29. V. Paulsen, *Syndetic sets, paving and the Feichtinger conjecture*, Proc. Amer. Math. Soc. **139** (2011), 1115–1120.
30. D. Plaumann, R. Sinn, D. Speyer, C. Vinzant, *Computing Hermitian determinantal representations of hyperbolic curves*, Internat. J. Algebra Comput. **25** (2015), no. 8, 1327–1336.
31. N. Weaver, *The Kadison–Singer problem in discrepancy theory*, Discrete Math. **278** (2004), no. 1–3, 227–239.

Department of Mathematics, University of Oregon, Eugene, OR 97403–1222, USA
E-mail address: mbownik@uoregon.edu

Department of Mathematics, University of Missouri-Columbia, Columbia, MO 65211, USA
E-mail address: casazzap@missouri.edu

Department of Mathematics, Princeton University, Princeton, NJ 08544–1000, USA
E-mail address: amarcus@princeton.edu

Department of Mathematics and Computer Science, Saint Louis University, 221 N. Grand Blvd., St. Louis, MO 63103, USA
E-mail address: speegled@slu.edu