MATRIX EXTREME POINTS AND FREE EXTREME POINTS OF FREE SPECTRAHEDRA

AIDAN EPPERLY, ERIC EVERT, J. WILLIAM HELTON, AND IGOR KLEP

Abstract. Free spectrahedra are dimension free solution sets to linear matrix inequalities of the form

$$L_A(X) = I_d \otimes I_n + A_1 \otimes X_1 + A_2 \otimes X_2 + \cdots + A_g \otimes X_g \succeq 0,$$

where the A_i and X_i are symmetric matrices and the X_i have any size $n \times n$. Free spectrahedra are ubiquitous in systems engineering, operator algebras, and the theory of matrix convex sets. Matrix and free extreme points of free spectrahedra are particularly important. We present theoretical, algorithmic, and experimental results illuminating basic properties of extreme points. For example, though many authors have studied matrix and free extreme points, it has until now been unknown if these two types of extreme points are actually different. This paper settles that issue.

We also present and analyze several algorithms. Namely, we perfect an algorithm for computing an expansion of an element of a free spectrahedron in terms of free extreme points. We also give algorithms for testing if a point is matrix extreme and for computing matrix extreme points that are not free extreme.

1. Introduction

Semidefinite programming [BPT12] is a generalization of linear programming which has played a profound role in applied mathematics. It is based on optimization of linear functionals over convex sets defined by linear matrix inequalities, namely, inequalities of the form

$$L_A(X) = I + A_1 X_1 + \cdots + A_g X_g \succeq 0.$$

Here the X_i are real numbers and the set of solutions is a convex set called a spectrahedron. Linear programming amounts to the very special case where every A_i is diagonal.

These inequalities also make sense when the X_i are symmetric matrices of any size, $n \times n$, and enter the formula though Kronecker's (tensor) product $A_i \otimes X_i$. The solution set of $L_A(X) \succeq 0$, denoted $\mathcal{D}_A$, is called a free spectrahedron since it contains matrices of all sizes and the defining "linear pencil" is "free" of the sizes of the matrices. The subset $\mathcal{D}_A(n)$ of $\mathcal{D}_A$ consisting of g-tuples of $n \times n$ matrices X_i is called the nth level of $\mathcal{D}_A$. Free spectrahedra are important examples of matrix convex sets [EW97, Kri19].

Extreme points are of great importance to understanding convex sets $\mathcal{C}$ for reasons including:

(a) the set of optimizers of a linear functional over $\mathcal{C}$ contains an extreme point;
(b) a "Caratheodory Krein-Milman type" theorem which expresses a given element of a matrix convex set as a convex combination of extreme points.

Free spectrahedra are level-wise convex, so there is the usual notion of scalar convex combination, but more powerful are matrix convex combinations; here contraction matrices summing to the identity play the role of the coefficients in the combination. These matrix convex combinations allow for *convex combinations of matrix tuples of different sizes* thereby leading to a close relationship between the various levels of a matrix convex set. In particular the geometry of the nth level of

Date: January 24, 2024.
2020 *Mathematics Subject Classification.* Primary 47L07, 13J30. Secondary 46L07, 90C22.
Key words and phrases. Spectrahedron, linear matrix inequality (LMI), matrix convex set, matrix extreme point, Arveson extreme point, free extreme point, dilation theory, Caratheodory expansion.
IK was supported by the Slovenian Research Agency grants J1-2453, J1-3004, N1-0217 and P1-0222.

a matrix convex set directly impacts all other levels of the set, thereby making matrix convexity much more restrictive than classical convexity.

Matrix convex sets have many natural notions of extreme points. Three of particular importance are *Euclidean, matrix* and *free extreme points* with their set containments being

$$\text{free extreme} \quad \subseteq \quad \text{matrix extreme} \quad \subseteq \quad \text{Euclidean extreme}$$

Here Euclidean refers to old-fashioned classical extreme points. This paper focuses on matrix and free extreme points and treats many aspects of them, both theoretical and numerical.

1.1. **The focus of this paper.** Surprisingly, despite recent advancements in the understanding of extreme points of matrix convex sets, it has remained unknown if there are matrix extreme points of a real free spectrahedron that are not free extreme. Though morally speaking such examples should exist as the definition of a matrix extreme point is evidently weaker than the definition of a free extreme point, for real free spectrahedra, this question has remained open.

For example, extensive experiments in [EFHY21] optimizing random linear functionals over random free spectrahedra found that while optimizers were with extremely high probability free extreme points the rest were merely euclidean extreme or (a few) were numerically bad.

At the time of running these experiments, we did not have the algorithmic tools to check if a given point is matrix extreme, however, out of several million trials, all but a few hundred of the points generated were Arveson extreme. This plus various tests of sufficient conditions led us to wonder if matrix but not Arveson extreme points exist.

In this paper we exhibit real free spectrahedra in three and in four variables with matrix extreme points that are not free extreme (no examples in two variables are known). A major difficulty is that this involves big matrices (e.g., 12×12) and the (eigenvalue-type) computations must be done algebraically. In another direction we prove that if the A_j are 2×2 matrices, then matrix and free extreme points of $\mathcal{D}_A$ are the same.

In addition we report systematic numerical experiments (for free spectrahedra in $g = 2, 3, 4$ variables) aimed at three goals:

(1) indicate if matrix extreme points which are not free extreme are rare,
(2) develop an effective algorithm that expands a point in a free spectrahedron as a matrix convex combination of free or matrix extreme points, i.e., find a free Caratheodory expansion,
(3) see how the observed size of a free Caratheodry expansion compares to the best theoretical upper bound.

Working on these two objectives together is natural since the underlying algorithms are related. Of independent interest is that to meet these objectives we develop a new technique called *Nullspace Purification*, which when added to the underlying algorithms greatly improves their accuracy.

The conclusions of our experiments are as follows. There is a strong indication that matrix but not free extreme points are not rare. In addition, the free Caratheodory expansion with Nullspace Purification had a low failure rate, especially for $g > 2$. Moreover, the addition of Nullspace purification to the previously best known approach for computing a free Caratheodory expansion [EFHY21] significantly improved the success rate. In addition, the "size" of the expansion was considerably lower than the provable worst case. This illustrates that free Caratheodory expansions in practice are typically even nicer than than theory suggests. See Section 8 for further comments on the observed patterns. For quantitative versions of the assertions just made, see the section on Conclusions, Section 9. Further observations on patterns are at the end of Section 7 and 8 and tables of data in these sections should be interesting to enthusiasts.

The experiments reported on in this article where run using NCSE (NonCommutative Spectrahedron Extreme) [EEd+21], an NCAlgebra [dOHMS17] package[1] for performing computations on elements of free spectrahedra. As part of our investigation, we added a number of new functions

[1]Both NCSE and NCAlgebra are written using Mathematica.

to NCSE. This includes algorithms for the aforementioned Nullspace Purification technique and algorithms for determining if a given element of a free spectrahedron is a matrix extreme point.

1.2. **Context and related work.** A major movement over the last few decades is to extend theorems involving polynomials, rational functions and power series expansions on $\mathbb{C}^g$ and $\mathbb{R}^g$ to noncommuting analogs acting on g tuples of matrices or operators leading to the development of the booming area of *free analysis* [BGM06, Pop06, dOHMP09, Voi10, KVV14, AM15, AJP20, JMS21]. Early in this direction was free probability which ultimately applied to random matrices of large size, cf. [MS17]. A later avenue is free real algebraic geometry which extends real algebraic geometry from $\mathbb{R}^g$ to g-tuples of matrices, [HM12, NT15, BKP16, Kri19].

Another major direction in free analysis involves Semidefinite programming (SDP) and LMIs as we are studying here. Related structures whose techniques are very useful for free LMIs have been studied by pure mathematicians for decades in the area of operator theory. Extreme points in this context, often called Arveson boundary points, go back to Arveson's seminal work [Arv69] and have been studied by many authors, e.g., see [Ham79, Agl88, MS98, DM05, Arv08, Kle14, DK15, FHL18, Pas22]. The overt introduction of matrix extreme points was in [WW99] with many advances at the same time due to Farenick [Far00, Far04]. Recently various works used free spectrahedra to study spectrahedral inclusion [HKM13, DDOSS17, FNT17, Zal17, HKMS19], and profound implications for quantum information theory were exposited in [BN18, BJN22]. A booming area based on "noncommutative optimization" [DLTW08, PNA10, BKP16, MBM21, WM21] is quantum game theory [GdLL19, DlCDN20, PR21]. Free LMIs can also model convex structures occurring in linear control and systems engineering problems specified entirely by signal flow diagrams, cf. [HMPV09].

The material in this paper partially appears in the undergraduate honor's thesis of Aidan Epperly at UC San Diego under the direction of J. William Helton. The remainder of the introduction turns to details, giving precise definitions and background theorems.

1.3. **Matrix notation.** Let $M_{m\times n}(\mathbb{R})^g$ denote the set of g-tuples of $m\times n$ matrices with real entries, and $M_n(\mathbb{R}) = M_{n\times n}(\mathbb{R})$. Similarly, let $SM_n(\mathbb{R})^g$ denote the set of all g-tuples of real symmetric $n \times n$ matrices and set $SM(\mathbb{R})^g = \cup_n SM_n(\mathbb{R})^g$. A matrix $B \in M_n(\mathbb{R})$ is **positive semidefinite** (PSD), denoted $B \succeq 0$, if it is symmetric and all of its eigenvalues are nonnegative. Given two symmetric matrices $B_1, B_2 \in M_n(\mathbb{R})$, let $B_1 \succeq B_2$ denote that $B_1 - B_2$ is PSD.

1.4. **Free spectrahedra and linear matrix inequalities.** In this paper, we primarily concern ourselves with a specific class of convex sets called free spectrahedra. A free spectrahedron is a matrix convex set that can be defined by a linear matrix inequality. Fix a tuple $A \in SM_d(\mathbb{R})^g$ of $d \times d$ symmetric matrices. A **monic linear pencil** $L_A(x)$ is a sum of the form

$$L_A(x) = I_d + A_1x_1 + A_2x_2 + \cdots + A_gx_g.$$

Given a tuple $X \in SM_n(\mathbb{R})^g$, the **evaluation** of L_A at X is

$$L_A(X) = I_d \otimes I_n + A_1 \otimes X_1 + A_2 \otimes X_2 + \cdots + A_g \otimes X_g$$

where $\otimes$ denotes the Kronecker Product. A **linear matrix inequality** is an inequality of the form $L_A(X) \succeq 0$. Let $\Lambda_A(X)$ denote the homogeneous linear part of $L_A(X)$, i.e.,

$$\Lambda_A(X) = A_1 \otimes X_1 + A_2 \otimes X_2 + \cdots + A_g \otimes X_g.$$

Given a g-tuple $A \in SM_n(\mathbb{R})^g$ and a positive integer n, we define the set $\mathcal{D}_A(n)$ by

$$\mathcal{D}_A(n) := \{X \in SM_n(\mathbb{R})^g : L_A(X) \succeq 0\}.$$

$\mathcal{D}_A(n)$ is called **free spectrahedron $\mathcal{D}_A$ at level n** Define the **free spectrahedron $\mathcal{D}_A$** to be the union over all n of the free spectrahedron $\mathcal{D}_A$ at level n, i.e.

$$\mathcal{D}_A := \bigcup_{n=1}^{\infty} \mathcal{D}_A(n) \subseteq SM(\mathbb{R})^g.$$

We say a free spectrahedron is **bounded** if there is some real number C so that

$$CI_n - \sum_{i=1}^{g} X_i^2 \succeq 0$$

for all $X = (X_1, X_2, \ldots, X_g) \in \mathcal{D}_A(n)$ and all positive integers n. It is routine to show that a free spectrahedron is bounded if and only if $\mathcal{D}_A(1)$ is bounded [HKM13]. In our definition of a free spectrahedron, we use a non-strict inequality. All free spectrahedra defined in this way are closed in the sense that each $\mathcal{D}_A(n)$ is closed.

1.4.1. *Minimal Defining Tuples.* Throughout this paper, the size of the defining tuple of a free spectrahedron will play an important role in our analysis. To make the size of the defining tuple of a free spectrahedron well-defined, we it is important to consider **minimal defining tuples**. Using [HKM13], we define a minimal defining tuple $\widetilde{A} \in SM_d(\mathbb{R})^g$ of a free spectrahedron $\mathcal{D}_A$ as a tuple of minimal size such that $\mathcal{D}_{\widetilde{A}} = \mathcal{D}_A$. That is to say if $\widetilde{A} \in SM_d(\mathbb{R})^g$ is a minimal defining tuple of $\mathcal{D}_A$ and $\widehat{A} \in SM_n(\mathbb{R})^g$ for $n < d$, then $\mathcal{D}_{\widehat{A}} \neq \mathcal{D}_{\widetilde{A}}$. The minimal defining tuple $\widetilde{A}$ is unique up to unitary equivalence [HKM13, Theorem 3.12 and Corollary 3.18]. Here, two tuples $X, Z \in \Gamma(n)$ are said to be **unitarily equivalent** if there exists some unitary matrix U such that $X_j = U^T Z_j U$. Throughout the paper, we will only consider minimal defining tuples.

1.4.2. *Homogeneous free spectrahedra.* Our proofs make heavy use of homogeneous free spectrahedra and of a canonical association between free spectrahedra and homogeneous free spectrahedra. Given a tuple $(A_0, \ldots, A_g) = (A_0, A) \in SM_d(\mathbb{R})^{g+1}$ the **homogeneous free spectrahedron $\mathcal{H}_{(A_0,A)}$ at level** n, denoted $\mathcal{H}_{(A_0,A)}(n)$, is given by

$$\mathcal{H}_{(A_0,A)}(n) := \left\{(X_0, \ldots, X_g) \in SM_n(\mathbb{R})^{g+1} : \Lambda_{(A_0,A)}(X_0, X) \succeq 0\right\}$$

and the **homogeneous free spectrahedron** $\mathcal{H}_{(A_0,A)}$ is defined by

$$\mathcal{H}_{(A_0,A)} = \cup_{n=1}^{\infty} \mathcal{H}_{(A_0,A)}(n) \subseteq SM(\mathbb{R})^{g+1}$$

We write the tuples $(X_0, \ldots, X_g)$ and $(A_0, \ldots, A_g)$ as (X_0, X) and (A_0, A), respectively, since we will frequently pass from nonhomogeneous free spectrahedra $\mathcal{D}_A$ with elements $X \in SM(\mathbb{R})^g$ to homogeneous free spectrahedra $\mathcal{H}_{(I,A)}$ with elements $(X_0, X) \in SM(\mathbb{R})^{g+1}$. We call X_0 and A_0 the **inhomogeneous component** of the corresponding tuple.

Given a free spectrahedron $\mathcal{D}_A$ in g variables, the **homogenization** of $\mathcal{D}_A$, denoted $\mathfrak{H}(\mathcal{D}_A)$, is the homogeneous free spectrahedron

$$\mathfrak{H}(\mathcal{D}_A) := \mathcal{H}_{(I,\check{A})}$$

where $\check{A}$ is any minimal defining tuple for $\mathcal{D}_A$. Here we emphasize the minimality of $\check{A}$, as it is required for the homogenization to be well-defined, see [Eve21]. On the other hand, given a homogeneous free spectrahedron $\mathcal{H}_{(A_0,A)}$, let $\mathfrak{H}^{-1}(\mathcal{H}_{(A_0,A)})$ denote the set

$$\mathfrak{H}^{-1}(\mathcal{H}_{(A_0,A)}) = \{X \in SM(\mathbb{R})^g |\ (I, X) \in \mathcal{H}_{(A_0,A)}\}.$$

A homogeneous free spectrahedron $\mathcal{H}_{(A_0,A)}$ is **sectionally bounded** if $\mathfrak{H}^{-1}(\mathcal{H}_{(A_0,A)})$ is bounded. We say that $\mathcal{H}_{(A_0,A)}$ is a **positive homogeneous free spectrahedron** if it contains $(1, 0, \ldots, 0) \in SM_1(\mathbb{R})^{g+1}$ in the interior of $\mathcal{H}_{(A_0,A)}(1)$. Note that the homogenization $\mathfrak{H}(\mathcal{D}_A)$ of any free spectrahedron $\mathcal{D}_A$ is always a positive homogeneous free spectrahedron.

We extend $\mathfrak{H}$ and $\mathfrak{H}^{-1}$ to matrix tuples as follows. Given a tuple $X \in SM(\mathbb{R})^g$ define

$$\mathfrak{H}(X) = (I, X) \in SM(\mathbb{R})^{g+1}.$$

On the other hand, given a matrix tuple $(X_0, X) \in SM(\mathbb{R})^{g+1}$ with $X_0 \succeq 0$, define

$$\mathfrak{H}^{-1}(X) = X_0^{\dagger/2} X X_0^{\dagger/2} \in SM(\mathbb{R})^g.$$

Here $X_0^{\dagger/2}$ denotes the PSD square root of the Moore-Penrose pseudoinverse of X_0. We refer the reader to [Eve21, Kri19] for further discussion of homogeneous free spectrahedra.

1.5. **Matrix Convex Sets.** Given some finite collection $\{X^i\}_{i=1}^{\ell}$ with $X^i \in SM_{n_i}(\mathbb{R})^g$ for each $i = 1, 2, \dots, \ell$, a **matrix convex combination** of $\{X^i\}_{i=1}^{\ell}$ is a sum of the form

$$\sum_{i=1}^{\ell} V_i^T X^i V_i \qquad \text{with} \qquad \sum_{i=1}^{\ell} V_i^T V_i = I_n$$

where $V_i \in M_{n_i \times n}(\mathbb{R})$ and

$$V_i^T X^i V_i = \left(V_i^T X_1^i V_i, V_i^T X_2^i V_i, \dots, V_i^T X_g^i V_i\right) \in SM_n(\mathbb{R})^g$$

for all $i = 1, 2, \dots, \ell$. We emphasize that the tuples X^i need not be the same size. A matrix convex combination is called **proper** if V_i is surjective for all $i = 1, 2, \dots, \ell$. A set $\Gamma \subseteq SM(\mathbb{R})^g$ is **matrix convex** if it is closed under matrix convex combinations. The **matrix convex hull** of a set $\Gamma \subseteq SM(\mathbb{R})^g$ is the set of all matrix convex combinations of the elements of Γ. It is straightforward to show that free spectrahedra are matrix convex.

A set $\Gamma \subseteq SM(\mathbb{R})^g$ is a **matrix cone** if given any finite collection $\{X^i\}_{i=1}^{\ell}$ with $X^i \in SM_{n_i}(\mathbb{R})^g$ for each $i = 1, 2, \dots, \ell$ and $V_{n_i} \in M_{n_i \times n}(\mathbb{R})$, then

$$\sum_{i=1}^{\ell} V_i^T X^i V_i \in \Gamma.$$

1.6. **Extreme points of matrix convex sets.** The extreme points of matrix convex sets and free spectrahedra are of particular interest, since they have Krein-Milman type spanning properties [WW99, Kri19, EH19]. The paper will primarily consider three types of extreme points: Euclidean extreme points, matrix extreme points, and free extreme points.

Given a matrix convex set Γ, we say $X \in \Gamma(n)$ is a **Euclidean extreme point** (EEP) of Γ if X cannot be written as a nontrivial classical convex combination of points in $\Gamma(n)$. This is the same as being a classical extreme point of $\Gamma(n)$. We let $\partial^{\text{Euc}}\Gamma$ denote the set of all the EEPs of Γ.

We say a point $X \in \Gamma(n)$ is a **matrix extreme point** (MEP) of Γ if whenever X is written as a proper matrix convex combination

$$X = \sum_{i=1}^{\ell} V_i^T X^i V_i \qquad \text{with} \qquad \sum_{i=1}^{\ell} V_i^T V_i = I_n$$

of points $X^i \in \Gamma$ for $i = 1, 2, \dots, \ell$, then for every $i = 1, 2, \dots, \ell$, we have $V_i \in M_n(\mathbb{R})$ and X is unitarily equivalent to X^i. We let $\partial^{\text{mat}}\Gamma$ denote the set of all the MEPs of Γ.

Finally, we say a point $X \in \Gamma(n)$ is a **free extreme point** (FEP) of Γ if whenever X is written as a matrix convex combination

$$X = \sum_{i=1}^{\ell} V_i^T X^i V_i \qquad \text{with} \quad \sum_{i=1}^{\ell} V_i^T V_i = I_n$$

of points $X^i \in \Gamma$ with $V_i \neq 0$ for each i, then for all $i = 1, 2, \dots, \ell$ either $V_i \in M_n(\mathbb{R})$ and X is unitarily equivalent to X^i or $V_i \in M_{n_i \times n}(\mathbb{R})$ where $n_i > n$ and there exists $Z^i \in \Gamma$ such that $X \oplus Z^i$ is unitarily equivalent to X^i. We let $\partial^{\text{free}}\Gamma$ denote the set of all the FEPs of Γ.

1.6.1. *Irreducible matrix tuples.* Given a matrix $M \in M_n(\mathbb{R})$, a subspace $N \subseteq \mathbb{R}^n$ is a **reducing subspace** if both N and $N^\perp$ are invariant subspaces of M, which is to say that N is a reducing subspace of M if $MN \subseteq N$ and $MN^\perp \subseteq N^\perp$. A tuple $X \in SM_n(\mathbb{R})^g$ is **irreducible** (over $\mathbb{R}$) if the matrices $X_1, \dots, X_g$ have no common reducing subspaces in $\mathbb{R}^n$; a tuple is **reducible** (over $\mathbb{R}$) if it is not irreducible. Since we will always be working over $\mathbb{R}$ in this paper, we will drop the use of "over $\mathbb{R}$" and simply refer to tuples as reducible or irreducible.

Remark 1.1. *Given a matrix convex set Γ, if $X \in \Gamma$ is a MEP, then X is irreducible.*

1.7. **Significance of FEPs and MEPs.** The concept of MEPs was introduced by Webster and Winkler in [WW99]. It is known that the matrix convex hull of the MEPs of a closed bounded matrix convex set Γ is equal to Γ [Kri19]. In particular, this means that every point in a free spectrahedron can be written as a matrix convex combination of MEPs. In [EH19] it was shown that every bounded free spectrahedron $\mathcal{D}_A$ is the matrix convex hull of its FEPs. Moreover, the FEPs are the smallest set with this property. We call an expansion of an element of a free spectrahedron in terms of FEPs a **free Caratheodory expansion.**

Theorem 1.2 ([EH19, Theorem 1.1])**.** *Let $A \in SM_d(\mathbb{R})^g$ such that $\mathcal{D}_A$ is a bounded free spectrahedron. Then $\mathcal{D}_A$ is the matrix convex hull of its FEPs. Furthermore, if $E \subseteq \mathcal{D}_A$ is a set of irreducible tuples which is closed under unitary equivalence and whose matrix convex hull is equal to $\mathcal{D}_A$, then E must contain the FEPs of $\mathcal{D}_A$.*

Thus, the FEPs are, in some sense, the correct notion of extreme points for free spectrahedra, that is, the point X in the free spectrahedron $\mathcal{D}_A$ has a free Caratheodory expansion. There are some differences when one considers these notions over $\mathbb{C}$ instead of over $\mathbb{R}$, cf. [Pas22].

1.8. **Guide to Main Results.** The paper is organized as follows. Section 2 focuses on the background theory underpinning our results. Of note is Theorem 2.2, a result of Kriel [Kri19], which allows us to characterize MEPs. Section 3 proves that the FEPs and MEPs of $\mathcal{D}_A$ are the same when $A \in SM_2(\mathbb{R})^g$. In Section 4 we exhibit examples of tuples that are MEPs but not FEPs. Section 5 and Section 6 describe algorithms for generating exact and numerical extreme points respectively. Section 6 also describes an algorithm for computing the free Caratheodory expansion of a given $X^0 \in \mathcal{D}_A$, as well as a technique called Nullspace Purification which greatly improves accuracy of our experiments. Section 7 focuses heavily on empirical observations resulting from a number of numerical experiments; these experiments lead us to believe that non-free matrix extreme points are not rarities. Section 8 tests our free Caratheodory algorithm together with Nullspace Purification and establishes that it is successful. Also here we see that the observed size of a free Caratheodory expansion is statistically much smaller than the theoretical upper bound. Section 9 summarizes our findings.

2. Background Theory

2.1. **Characterizing Extreme Points.** An important class of extreme points that has yet to be discussed is the Arveson extreme points which arise from dilation theory. Given an $X \in SM_n(\mathbb{R})^g$, say $Y \in SM_{n+k}(\mathbb{R})^g$ is a **dilation**, or more specifically a **k-dilation**, of X if Y is of the form

$$Y = \begin{pmatrix} X & \beta \\ \beta^T & \gamma \end{pmatrix}$$

for $\beta \in M_{n\times k}(\mathbb{R})^g$ and $\gamma \in SM_k(\mathbb{R})^g$. A point $X \in \Gamma$ is an **Arveson extreme point** (AEP) of Γ if $Y \in \Gamma$ being a 1-dilation of X implies that $\beta = 0$. We can provide a similar characterization of a EEP in the language of dilation theory:

Proposition 2.1 ([EHKM18, Corollary 2.3]). *A point $X \in SM_n(\mathbb{R})^g$ in a free spectrahedron $\mathcal{D}_A$ is a EEP of $\mathcal{D}_A$ if and only if*

$$Y = \begin{pmatrix} X & \beta \\ \beta & \gamma \end{pmatrix} \in \mathcal{D}_A$$

implies $\beta = 0$ for $\beta \in SM_n(\mathbb{R})^g$.

A theorem of Kriel gives a similarly flavored result regarding MEPs. A simple proof is given in an online appendix for the sake of completeness.

Theorem 2.2 ([Kri19, Theorem 6.5.c]). *A point X in a bounded free spectrahedron $\mathcal{D}_A$ is a MEP of $\mathcal{D}_A$ if and only if (I, X) is on a classical extreme ray of $\mathcal{H}_{(I,A)}$.*

Corollary 2.3. *For a point X in a bounded free spectrahedron $\mathcal{D}_A$ the following are equivalent:*

(i) *X is a MEP of $\mathcal{D}_A$;*
(ii) *(I, X) is on a classical extreme ray of $\mathcal{H}_{(I,A)}$;*
(iii) *$(\beta_0, \beta) \in SM_n(\mathbb{R})^{g+1}$ and*

$$\ker \Lambda_{(I,A)}(I, X) \subseteq \ker \Lambda_{(I,A)}((\beta_0, \beta)) \implies (\beta_0, \beta) \in \operatorname{span}((I, X));$$

(iv) *For $(\beta_0, \beta), (\gamma_0, \gamma) \in SM_n(\mathbb{R})^{g+1}$,*

$$(Y_0, Y) = \begin{pmatrix} (I, X) & (\beta_0, \beta) \\ (\beta_0, \beta) & (\gamma_0, \gamma) \end{pmatrix} \in \mathcal{H}_{(I,A)}$$

implies $(\beta_0, \beta) \in \operatorname{span}((I, X))$.

Proof. The equivalence of Item (i) and Item (ii) follows from Theorem 2.2. The equivalence of Item (ii) and Item (iii) is [RG95, Corollary 4] by viewing (I, X) as an element of $\mathbb{R}^{\frac{n(n+1)(g+1)}{2}}$. The equivalence of Item (iii) and Item (iv) follows the same argument as the proof of [EHKM18, Corollary 2.3]. Details are given in the online appendix. ∎

Using these characterizations of EEPs and MEPs, we arrive at the following known result, see [EHKM18, Theorem 1.1].

Proposition 2.4. *Let $\mathcal{D}_A$ be a bounded free spectrahedron.*

(1) *A tuple X is a FEP of $\mathcal{D}_A$ if and only if X is an irreducible AEP of $\mathcal{D}_A$.*
(2) *If a tuple X is a FEP of $\mathcal{D}_A$, then X is a MEP of $\mathcal{D}_A$.*
(3) *If a tuple X is a MEP of $\mathcal{D}_A$, then X is a EEP of $\mathcal{D}_A$.*
(4) *If a tuple X is an AEP of $\mathcal{D}_A$, then X is a EEP of $\mathcal{D}_A$.*

Proof. Item (1) and Item (4) are the subject of [EHKM18, Theorem 1.1], where the proof is given working over $\mathbb{C}$. The proof of Item (4) can be used over $\mathbb{R}$ without modification, and the proof of Item (1) over $\mathbb{R}$ is given by [EH19, Theorem 1.2]. Item (2) follows from [EH19, Theorem 1.1] which is given as Theorem 1.2 here. Item (3) follows from the fact that if X can be written as a nontrivial classical convex combination $X = \alpha_1 X^1 + \alpha_2 X^2 + \cdots + \alpha_\ell X^\ell$, then $(I, X) = \alpha_1(I, X^1) + \alpha_2(I, X^2) + \cdots + \alpha_\ell(I, X^\ell)$ is a nontrivial convex combination of points in $\mathcal{H}_{(I,A)}$. ∎

Corollary 2.5. *Let $\mathcal{D}_A$ be a bounded free spectrahedron. If X is a MEP of $\mathcal{D}_A$, then X is a FEP of $\mathcal{D}_A$ if and only if X is an AEP of $\mathcal{D}_A$.*

Proof. Straightforward. ∎

Thus, if X is a MEP of $\mathcal{D}_A$, to show that X is not a FEP it is sufficient to show it is not an AEP. Checking if a tuple is an AEP is equivalent to solving the upcoming linear system Equation (2.1).

2.1.1. *Extreme points and linear systems.* One can determine if an element of a free spectrahedron is extreme by solving a linear system. Given a free spectrahedron $\mathcal{D}_A$ and a point $X \in \mathcal{D}_A$ we let

$$k_{A,X} := \dim \ker L_A(X)$$

and let $K_{A,X}$ be an $nd \times k_{A,X}$ matrix whose columns form an orthonormal basis for $\ker L_A(X)$.

Theorem 2.6. *Let $\mathcal{D}_A$ be a bounded free spectrahedron and $X \in \mathcal{D}_A(n)$.*

(1) *X is an AEP of $\mathcal{D}_A$ if and only if the only solution to the homogeneous linear equations*

$$\Lambda_A(\beta^T)K_{A,X} = (A_1 \otimes \beta_1^T + \cdots + A_n \otimes \beta_g^T)K_{A,X} = 0 \tag{2.1}$$

in the unknown $\beta \in M_{n\times 1}(\mathbb{R})^g$ is $\beta = 0$.

(2) *X is a EEP of $\mathcal{D}_A$ if and only if the only solution to the homogeneous linear equations*

$$\Lambda_A(\beta)K_{A,X} = (A_1 \otimes \beta_1 + \cdots + A_n \otimes \beta_g)K_{A,X} = 0 \tag{2.2}$$

in the unknown $\beta \in SM_n(\mathbb{R})^g$ is $\beta = 0$.

(3) *X is a MEP of $\mathcal{D}_A$ if and only if the only solution to the homogeneous linear equations*

$$\Lambda_{(I,A)}(\beta_0, \beta)K_{A,X} = (I \otimes \beta_0 + A_1 \otimes \beta_1 + \cdots + A_n \otimes \beta_g)K_{A,X} = 0 \tag{2.3}$$

$$\langle (I, X), (\beta_0, \beta)\rangle = \operatorname{tr}(\beta_0 + X_1^T\beta_1 + \cdots + X_g^T\beta_g) = 0 \tag{2.4}$$

in the unknown $(\beta_0, \beta) \in SM_n(\mathbb{R})^{g+1}$ is $\beta = 0$.

Proof. The proofs for each of the above results are all quite similar, so details of the proofs for Item (1) and Item (2) will be omitted. The assertion given in Item (1) regarding AEPs is the content of [EH19, Lemma 2.1 (3)]. The assertion given in Item (2) regarding EEPs is the content of [EHKM18, Corollary 2.3]. It follows from [RG95, Corollary 3].

To show Item (3), we note that by Theorem 2.2 and Corollary 2.3, X is MEP if and only if for $(\beta_0, \beta), (\gamma_0, \gamma) \in SM_n(\mathbb{R})^{g+1}$

$$\Lambda_{(I,A)}(Y_0, Y) = \Lambda_{(I,A)}\left(\begin{pmatrix} (I, X) & (\beta_0, \beta) \\ (\beta_0, \beta) & (\gamma_0, \gamma)\end{pmatrix}\right) \succeq 0$$

implies $(\beta_0, \beta) \in \operatorname{span}((I, X))$.

Clearly $\ker \Lambda_{(I,A)}((I, X)) = \ker L_A(X)$ so $K_{A,X}$ is a matrix whose columns form an orthonormal basis of $\ker \Lambda_{(I,A)}((I, X))$ as well. If there is a (β_0, β) such that $(Y_0, Y) \in \mathcal{H}_{(I,A)}$ and (β_0, β) is not in the span of (I, X), then by applying the canonical shuffle to the above LMI and considering the appropriate Schur compliment, it follows that

$$\ker \Lambda_{(I,A)}((I, X)) \subseteq \ker \Lambda_{(I,A)}((\beta_0, \beta)).$$

Write $(\beta_0, \beta) = (\beta_0', \beta') + \alpha(I, X)$ for $\alpha \in \mathbb{R}$ and $(\beta_0', \beta') \neq 0$ satisfying Equation (2.4). Then

$$\ker \Lambda_{(I,A)}((I, X)) \subseteq \ker \Lambda_{(I,A)}((\beta_0, \beta)) = \ker(\Lambda_{(I,A)}(\alpha(I, X)) + \Lambda_{(I,A)}((\beta_0', \beta')))$$

implies that

$$\ker \Lambda_{(I,A)}((I, X)) \subseteq \ker \Lambda_{(I,A)}((\beta_0', \beta')).$$

Hence, there is a nonzero (β_0', β') satisfying Equation (2.3) and Equation (2.4).

Conversely, if there is (β_0, β) satisfying Equation (2.3) and Equation (2.4), then by taking $(\gamma_0, \gamma) = (I, 0)$ the argument above reverses to show X is not a MEP. ∎

Corollary 2.7. *Let $A \in SM_d(\mathbb{R})^g$ be a minimal defining tuple for $\mathcal{D}_A$ and let $X \in \mathcal{D}_A(n)$. Recall that $k_{A,X} = \dim \ker L_A(X)$. Then,*

(1) *if X is an AEP then*

$$gn \leq dk_{A,X}; \tag{2.5}$$

(2) *if X is a EEP extreme then*

$$\frac{g(n+1)}{2} \leq dk_{A,X}; \tag{2.6}$$

(3) *if X is a MEP then*

$$\frac{n(n+1)(g+1)}{2} \leq dnk_{A,X} + 1. \tag{2.7}$$

For brevity, we will often refer to

(1) $\left\lceil \frac{gn}{d} \right\rceil$ as the Arveson rank-nullity count;
(2) $\left\lceil \frac{g(n+1)}{2d} \right\rceil$ as the Euclidean rank-nullity count;
(3) $\left\lceil \frac{(n+1)(g+1)}{2d} - \frac{1}{nd} \right\rceil$ as the Matrix extreme rank-nullity count.

In this terminology, for X to be a certain type of extreme point of $\mathcal{D}_A$, the kernel dimension $k_{A,X}$ must be at least as large as the corresponding rank-nullity count.

Proof. The above inequalities result from comparing number of equations to number of unknowns in the homogeneous linear equations given in Theorem 2.6 and follow routine arguments. An example proof for Item (3) is given in the online appendix. ∎

2.1.2. *The Dilation Subspace.* The space of all solutions, β, to Equation (2.1) is useful to consider. This subspace of $M_{n\times 1}(\mathbb{R})^g$ shall be referred to as the **dilation subspace** of X with respect to $\mathcal{D}_A$. We call the dimension of this space the **dilation subspace dimension** of X with respect to $\mathcal{D}_A$, often abbreviated *dilDim*. The reference to $\mathcal{D}_A$ is often dropped when context makes it clear.

3. MEPs are always FEPs when $d = 2$

In this section we show that MEPs and FEPs coinincide when $d = 2$.

Theorem 3.1. *Let $A \in SM_2(\mathbb{R})^g$. Then $\partial^{\text{mat}}(\mathcal{D}_A) = \partial^{\text{free}}(\mathcal{D}_A)$.*

Proof. See Section 3.2. ∎

To prove Theorem 3.1, we separately consider the cases $g = 2$ and $g \geq 3$. The $g = 2$ case is handled using projective maps to show that any bounded free spectrahedron $\mathcal{D}_A$ with $A \in SM_2(\mathbb{R}^2)$ can be mapped to a canonical free spectrahedron, see Proposition 3.4. Furthermore we show in the upcoming Theorem 3.3 that MEPs are preserved under invertible projective transformations. The $g \geq 3$ case is handled by the following proposition which fully classifies all matrix and FEPs of free spectrahedra that satisfy $g \geq d(d+1)/2$.

Proposition 3.2. *Let $A \in SM_d(\mathbb{R})^g$. If $g \geq d(d+1)/2$, then $\mathcal{D}_A$ is not bounded and is not the matrix convex hull of its FEPs. Furthermore, such free spectrahedra either have exactly one MEP which is also a FEP in which case $g = d(d+1)/2$ or they have no extreme points at all.*

Proof. First note that $\dim(SM_d(\mathbb{R})) = d(d+1)/2$. If $g > d(d+1)/2$, then there exist constants $\alpha_1, \ldots, \alpha_g \in \mathbb{R}$ such that $\sum_{i=1}^g \alpha_i A_i \otimes I = 0$. As a consequence, for any $X \in \mathcal{D}_A$ and any $c \in \mathbb{R}$ we have $X + c(\alpha_1 I, \ldots, \alpha_g I) \in \mathcal{D}_A$. We conclude that $\mathcal{D}_A$ has no extreme points at all if $g > d(d+1)/2$.

Now suppose $g = d(d+1)/2$. If $\{A_1, \ldots, A_g\}$ is a linearly dependent set, then the above argument shows $\mathcal{D}_A$ has no extreme points. If $\{A_1, \ldots, A_g\}$ is a linearly independent set, then there exist constants $\alpha_1, \ldots, \alpha_g$ such that $\sum_{i=1}^g \alpha_i A_i = -I_2$. We will show that $(\alpha_1, \ldots, \alpha_g) \in \mathcal{D}_A(1)$ is the only MEP of $\mathcal{D}_A$. To this end let $X \in \mathcal{D}_A(n)$ and observe that for any $c \geq 0$ we have

$$L_A((1-c)(\alpha_1 I, \ldots, \alpha_g I) + cX) = I + (1-c)I \otimes -I + c\Lambda(X) = cL_A(X) \succeq 0.$$

It follows that $(1-c)(\alpha_1 I, \dots, \alpha_g I) + cX \in \mathcal{D}_A$ for all $c \geq 0$ hence X is not a EEP of $\mathcal{D}_A$ unless $X = (\alpha_1 I, \dots, \alpha_g I)$. Moreover, since MEPs are irreducible tuples, X is not a MEP of $\mathcal{D}_A$ unless $X = (\alpha_1, \dots, \alpha_g) \in \mathcal{D}_A(1)$. It then follows from [EHKM18, Proposition 6.1] that this tuple is both a MEP and FEP of $\mathcal{D}_A$. Finally, since $\mathcal{D}_A$ is unbounded but only has one FEP, it is clear that $\mathcal{D}_A$ cannot be the matrix convex hull of its FEPs. ∎

3.1. **Projective maps of free spectrahedra.** The proof of the $g = 2$ case of Theorem 3.1 makes heavy use of projective mappings of free spectrahedra. We introduce here the basic definitions and notation related to projective maps. The definitions we use are the same as those found in [Eve21], and we direct the reader there for a detailed discussion of projective maps.

Given a matrix $W \in M_{g+1}(\mathbb{R})$, we define a **linear transformation** $\mathcal{T}_W$ on $SM(\mathbb{R})^{g+1}$ by

$$\mathcal{T}_W(X_0, X) = \Big(\sum_{j=1}^{g+1} W_{1,j} X_{j-1}, \dots, \sum_{j=1}^{g+1} W_{g+1,j} X_{j-1} \Big) \qquad \text{for all } (X_0, X) \in SM(\mathbb{R})^{g+1}.$$

Here $W_{i,j}$ denotes the (i,j)th entry of W. Note that $\mathcal{T}_W$ is in fact a **free linear transformation** meaning that it respects direct sums and simultaneous unitary conjugation. That is, for tuples $(X_0, X) \in SM_n(\mathbb{R})^{g+1}$ and $(Y_0, Y) \in SM_m(\mathbb{R})^{g+1}$, we have

$$\mathcal{T}_W((X_0, X) \oplus (Y_0, Y)) = \mathcal{T}_W(X_0, X) \oplus \mathcal{T}_W(Y_0, Y).$$

Additionally, given a unitary matrix $U \in \mathbb{R}^{n \times n}$ we have

$$\mathcal{T}_W(U^T(X_0, X)U) = U^T \mathcal{T}_W(X_0, X)U.$$

Given a $g+1$ tuple $(A_0, A) \in SM_d(\mathbb{R})^{g+1}$, the image of $\mathcal{H}_{(A_0,A)}$ under $\mathcal{T}_W$ is the set

$$\mathcal{T}_W(\mathcal{H}_{(A_0,A)}) = \{\mathcal{T}_W(X_0, X) |\ (X_0, X) \in \mathcal{H}_{(A_0,A)}\}.$$

As one might expect, a linear transformation of a homogeneous free spectrahedron is again a homogeneous free spectrahedron, see [Eve21, Lemma 3.1]. If $\mathcal{H}_{(A_0,A)}$ and $\mathcal{T}_W(\mathcal{H}_{(A_0,A)})$ are both positive free spectrahedra, then we say that $\mathcal{T}_W$ is a **positive linear transformation** of $\mathcal{H}_{(A_0,A)}$.

For a tuple $A \in SM_d(\mathbb{R})^g$ and a matrix $W \in M_{g+1}(\mathbb{R})$ such that $\mathcal{T}_W$ is a positive linear transformation of $\mathfrak{H}(\mathcal{D}_A)$, we define the projective transformation $\mathcal{P}_W$ of $\mathcal{D}_A$ by

$$\mathcal{P}_W(\mathcal{D}_A) = \mathfrak{H}^{-1}(\mathcal{T}_W(\mathfrak{H}(\mathcal{D}_A))).$$

In the case that $\mathcal{P}_W(\mathcal{D}_A)$ is bounded, we define $\mathcal{P}_W$ on tuples $X \in \mathcal{D}_A$ by

$$\mathcal{P}_W(X) := \mathfrak{H}^{-1}(\mathcal{T}_W(\mathfrak{H}(X))$$

Note that this is well-defined as a consequence of [Eve21, Lemma 2.2].

This completes our required definitions. Before proceeding, we warn the reader that there are several idiosyncrasies which can in principle occur when working with projective maps in the noncommutative setting. Fortunately, we will not encounter these issues since we always consider bounded homogeneous free spectrahedra. See [Eve21] for a more detailed discussion.

We now show that MEPs are preserved under invertible projective transformations.

Theorem 3.3. *Let $\mathcal{D}_A$ and $\mathcal{D}_B$ be bounded free spectrahedra and suppose there exists some invertible matrix W such that $\mathcal{P}_W$ is an invertible projective map from $\mathcal{D}_A$ to $\mathcal{D}_B$. Then $X \in \mathcal{D}_A$ is a MEP of $\mathcal{D}_A$ if an only if $\mathcal{P}_W(X)$ is a MEP of $\mathcal{D}_B$.*

Proof. Using Theorem 2.2 we have that X is a MEP of $\mathcal{D}_A$ if and only if the only solutions to

$$\ker \Lambda_{(I,A)}(I, X) \subseteq \ker \Lambda_{(I,A)}(Y_0, Y)$$

satisfy $(Y_0, Y) = \alpha(I, X)$ for some $\alpha \in R$. Arguing by contrapositive, suppose X is not a MEP of $\mathcal{D}_A$. Then there is some tuple (Y_0, Y) which satisfies

$$\ker \Lambda_{(I,A)}(I, X) \subseteq \ker \Lambda_{(I,A)}(Y_0, Y)$$

and that there is no $\alpha \in R$ such that $(Y_0, Y) = \alpha(I, X)$. From this we find

$$\ker \Lambda_{\mathcal{T}_{W^{-T}}(I,A)}\left(\mathcal{T}_W(I,X)\right) \subseteq \ker \Lambda_{\mathcal{T}_{W^{-T}}(I,A)}\left(\mathcal{T}_W(Y_0,Y)\right).$$

Furthermore, we cannot have $\alpha\mathcal{T}_W(I,X) = \mathcal{T}_W(Y_0,Y)$ since $\mathcal{T}_W$ is an invertible linear transformation. As a consequence of [Eve21, Lemma 3.3] we have $\mathcal{T}_{W^{-T}}(I,A) = (I,B)$ from which it follows that $\mathcal{T}_W(I,X)$ is not on an extreme ray of $\mathcal{H}_{(I,B)}$. Moreover, since $\mathcal{D}_B$ is bounded by assumption, [Eve21, Lemma 2.2] shows that the inhomogeneous component of $\mathcal{T}_W(I,X)$ is PSD. It is then straightforward to show that (I,X) is not on an extreme ray of $\mathcal{H}_{(I,B)}$. We conclude that $\mathcal{P}_W(X)$ is not a MEP of $\mathcal{D}_B$, as claimed. ∎

Knowing that MEPs are preserved under invertible projective transformations, to treat the bounded $d = g = 2$ case in Theorem 3.1, it is sufficient to consider some canonical free spectrahedron which any other bounded $g = d = 2$ free spectrahedron can be projectively mapped onto. The canonical free spectrahedron we consider is the spin disk since its MEPs and FEPs are well understood, see [EHKM18, Proposition 7.5].

Proposition 3.4. *Let $A \in SM_2(\mathbb{R})^2$ and assume that the free spectrahedron $\mathcal{D}_A$ is bounded and let $\mathcal{D}_B$ be the spin disk. That is, $\mathcal{D}_B$ is the free spectrahedron with defining tuple*

$$B = \left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right).$$

Then there exists an invertible matrix W such that the map $X \mapsto \mathcal{P}_W(X)$ defined on $\mathcal{D}_A$ is a well-defined invertible projective transformation which maps $\mathcal{D}_A$ onto $\mathcal{D}_B$.

Proof. Write $A = (A_1, A_2)$ where

$$A_1 = \begin{pmatrix} a_{111} & a_{121} \\ a_{121} & a_{221} \end{pmatrix} \qquad A_2 = \begin{pmatrix} a_{112} & a_{122} \\ a_{122} & a_{222} \end{pmatrix}.$$

Define the matrix $W \in \mathbb{R}^{3\times 3}$ by

$$W = \begin{pmatrix} 1 & \frac{a_{111}+a_{221}}{2} & \frac{a_{112}+a_{222}}{2} \\ 0 & \frac{a_{111}-a_{221}}{2} & \frac{a_{112}-a_{222}}{2} \\ 0 & a_{121} & a_{122} \end{pmatrix}.$$

We will show that $\mathcal{P}_W$ is a well-defined invertible projective map from $\mathcal{D}_A$ onto $\mathcal{D}_B$ and that $\mathcal{P}_W$ has inverse $\mathcal{P}_{W^{-1}}$. To accomplish this we must show that W is invertible and that $\mathcal{T}_W$ is a positive linear transformation that maps $\mathcal{H}_{(I,A)}$ onto $\mathcal{H}_{(I,B)}$.

We first show that W is invertible. To this end note that if a_{122} and a_{121} are both equal to zero, then $\mathcal{D}_A$ is unbounded which is a contradiction. Next note that

$$\det(W) = (a_{111}a_{122} - a_{121}a_{112} + a_{121}a_{222} - a_{221}a_{122})/2.$$

Thus, if W is not invertible we have

$$a_{111}a_{122} - a_{121}a_{112} = a_{221}a_{122} - a_{121}a_{222},$$

from which it follows that the matrix

$$a_{122}A_1 - a_{121}A_2 = \begin{pmatrix} a_{111}a_{122} - a_{121}a_{112} & 0 \\ 0 & a_{221}a_{122} - a_{121}a_{222} \end{pmatrix}$$

is either positive or negative semidefinite. In either case, $\mathcal{D}_A$ is not bounded since the nonzero vector $\alpha(a_{122}, -a_{121})$ is then an element of $\mathcal{D}_A$ either for all $\alpha \geq 0$ or $\alpha \leq 0$, depending on whether the quantity $a_{111}a_{122} - a_{121}a_{112}$ is positive or negative. We conclude that W is invertible.

Next observe that a direct calculation shows $\mathcal{T}_{W^T}(I, B_1, B_2) = (I, A_1, A_2)$ hence $\mathcal{T}_{W^{-T}}(I, A_1, A_2) = (I, B_1, B_2)$. Using [Eve21, Lemma 3.1] then implies

$$\mathcal{T}_W(\mathcal{H}_{(I,A)}) = \mathcal{H}_{(I,B)}.$$

Since $\mathcal{H}_{(I,A)}$ and $\mathcal{H}_{(I,B)}$ are both positive homogeneous free spectrahedra, we obtain that $\mathcal{T}_W$ is a positive linear transformation from $\mathcal{H}_{(I,A)}$ to $\mathcal{H}_{(I,B)}$. Furthermore, $\mathcal{H}_{(I,B)}$ is sectionally bounded since $\mathcal{D}_B$ is bounded. It follows that the map $X \mapsto \mathcal{P}_W(X)$ defined on $\mathcal{D}_A$ is indeed a well-defined projective transformation that maps $\mathcal{D}_A$ into $\mathcal{D}_B$. The proof is completed by [Eve21, Lemma 3.3] and [Eve21, Lemma 3.4] which together show that $\mathcal{P}_W$ maps $\mathcal{D}_A$ onto $\mathcal{D}_B$ and that this map is invertible with inverse equal to $\mathcal{P}_{W^{-1}}$. ∎

3.2. **Proof of Theorem 3.1.** The following lemma is useful in the Proof of Theorem 3.1.

Lemma 3.5. *Let $A, B \in \mathbb{R}^{2\times 2}$ be linearly independent real symmetric matrices and assume A is positive semidefinite. Then there exists an $\alpha_0 \in \mathbb{R}$ such that $A + \alpha_0 B$ is rank one and positive semidefinite.*

Proof. The proof essentially follows from the fact that the set of PSD matrices is a closed pointed cone together with the fact that we are working with 2×2 matrices. In particular, observe that there exists some choice of $\beta \in \mathbb{R}$ such that $A+\beta B$ is not positive semidefinite. Now for each $\alpha \in \mathbb{R}$, let $\lambda_{\min}(A + \alpha B)$ denote the smallest eigenvalue of $A + \alpha B$. Then $\lambda_{\min}(A + \alpha B)$ is continuous in α, and $\lambda_{\min}(A)$ is nonnegative, and $\lambda_{\min}(A + \beta B)$ is negative. Therefore, there must exist some α_0 such that $\lambda_{\min}(A + \alpha_0 B) = 0$. Since A and B are linearly independent and symmetric, $A + \alpha_0 B$ is then a nonzero PSD rank deficient 2×2, hence it has rank 1. ∎

Proof of Theorem 3.1. The result in the case that $g \geq 3$ if proved as Proposition 3.2, so it is sufficient to consider $g \leq 2$. We first assume that $g = 2$ and that $\mathcal{D}_A$ is unbounded. Equivalently, assume that $\operatorname{span}(\{A_1, A_2\})$ contains a PSD matrix. We first argue that it is sufficient to consider tuples (A_1, A_2) of the form

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & b \end{pmatrix} \tag{3.1}$$

where $b \in \mathbb{R}$. To this end, note that if $cI \in \operatorname{span}(\{A_1, A_2\})$ for some $0 \neq c \in \mathbb{R}$ or if $\{A_1, A_2\}$ is linearly dependent, then we can repeat the argument used in Proposition 3.2 to conclude that all MEPs of $\mathcal{D}_A$ are also FEPs. Therefore we assume that $cI \notin \operatorname{span}(\{A_1, A_2\})$ and that $\{A_1, A_2\}$ is linearly independent. In this case, an application of Lemma 3.5 shows that $\operatorname{span}(\{A_1, A_2\})$ contains a rank one PSD matrix. Furthermore, it is straightforward to show that if there exists an invertible linear transformation on $SM(\mathbb{R})^2$ which maps A to B, then $\mathcal{D}_A$ contains a MEP which is not a FEP if and only if $\mathcal{D}_B$ contains a MEP that is not a FEP. Using this fact, we can without loss of generality assume that A_1 has rank 1 and that the spectrum of A_1 is $\{1, 0\}$. From here we can use the fact that unitarily equivalent tuples define the same free spectrahedron together with another invertible change of a variables to reduce to the case where A_1 and A_2 are as in equation (3.1).

Now, with (A_1, A_2) as above, we can use the Schur complement to conclude that $X \in \mathcal{D}_A$ iff

$$I + bX_2 \succeq 0 \qquad \text{and} \qquad P := I + X_1 - X_2(1 + bX_2)^\dagger X_2 \succeq 0.$$

Observe that if $X \in \mathcal{D}_A$ then

$$I + (X_1 + P) - X_2(1 + bX_2)^\dagger X_2 = 2P \succeq 0 \qquad \text{and} \qquad I + (X_1 - P) - X_2(1 + bX_2)^\dagger X_2 = 0$$

from which we obtain that $(X_1 + P, X_2) \in \mathcal{D}_A$ and $(X_1 - P, X_2) \in \mathcal{D}_A$. Moreover, if $P \neq 0$, then

$$(X_1 + P, X_2) \neq X \neq (X_1 - P, X_2).$$

That is, if X is a EEP of $\mathcal{D}_A$, then $I + X_1 - X_2(1 + bX_2)^\dagger X_2 = 0$.

However in this case we have

$$X_1X_2 = (X_2(1 + bX_2)^\dagger X_2 - I)X_2 = X_2(X_2(1 + bX_2)^\dagger X_2 - I) = X_2X_1.$$

We conclude that if $X \in \mathcal{D}_A(n)$ is a EEP of $\mathcal{D}_A$ and $n \geq 1$, then X is reducible. It follows from [EHKM18, Theorem 1.1] that all MEPs and FEPs of $\mathcal{D}_A$ are contained in $\mathcal{D}_A(1)$. From here, one can use [EHKM18, Proposition 6.1] to show that all MEPs of $\mathcal{D}_A$ are also FEPs.

Now suppose that $g = 2$ and that $\mathcal{D}_A$ is bounded. Using Proposition 3.4 shows that there exists some invertible projective transformation $\mathcal{P}_W$ which maps $\mathcal{D}_A$ onto $\mathcal{D}_B$ where

$$B = \left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right).$$

Furthermore using Theorem 3.3 shows that a tuple X is a MEP of $\mathcal{D}_A$ if and only if $\mathcal{P}_W(X)$ is a MEP of $\mathcal{D}_B$. Similarly, [Eve21, Theorem 3.7] show that X is a FEP of $\mathcal{D}_A$ if and only if $\mathcal{P}_W(X)$ is a FEP of $\mathcal{D}_B$. Combining this with [EHKM18, Proposition 7.5] which shows that every MEP of $\mathcal{D}_B$ is also a FEP completes the proof in this case.

Finally, for completeness, we mention that the proof when $g = 1$ is straightforward. In this case it is easy to show that all MEPs of $\mathcal{D}_A$ are found at level 1 of $\mathcal{D}_A$, hence all MEPs are FEPs. ∎

4. Exact non-free matrix extreme Points

In this section we present examples of non-free matrix extreme points for free spectrahedra when $g = 3$ and $g = 4$. For $g = 2$ the existence of non-free matrix extreme points is not known. To prove an example has our claimed properties, numerical (floating point) calculations do not suffice due to possible numerical errors. However, usually it is difficult to find extreme points of free spectrahedra with exact arithmetic. Even in the smallest nontrivial case where the defining tuple $A \in SM_2(\mathbb{R})^2$ and the desired extreme point $X \in \mathcal{D}_A(2)$, exactly computing X by optimizing a linear functional requires computing an exact arithmetic solution to a semidefinite program with six variables and the constraint $L_A(X) \succeq 0$ where $L_A(X)$ is 4×4.

If the size of the defining tuple or the extreme point is greater than two, then computing even just a boundary point would require finding an exact solution α for $\det(L_A(\alpha X)) = 0$, which is often not possible in radicals since $\det(L_A(\alpha X))$ is generally a polynomial of degree greater than five in α. As a consequence, it can in some cases be impossible to express boundary points, let alone extreme points, using radicals. Algorithms which sometimes yield exact extreme points (typically expressed using roots of some polynomial) are the subject of Section 5.

4.1. $g = 3$ non-free matrix extreme example.

Now we give an example for $g = 3$ of a bounded free spectrahedron $\mathcal{D}_A$ and a non-free matrix extreme point Y in it. In this case, A and Y have entries which are algebraic numbers and we have proved using exact arithmetic that Y is in $\mathcal{D}_A$ and is not an AEP. To prove that Y is not a MEP one only needs to check that the Matrix Extreme Equation (2.3) has no solution. This we proved via floating point arithmetic by checking that the appropriate matrix has smallest singular value equal to 0.0318244 while largest singular value is not very large (< 5), hence it has no nullspace. Alternately, one could prove this using exact arithmetic, which would be considerably slower in this case.

Now we state our example. Let

$$A = \left(\begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & -1 & 1 & 1 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & -1 & -1 \\ 1 & 1 & -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \right)$$

be the defining tuple of the free spectrahedron $\mathcal{D}_A$ and

$$Y = \left(\begin{pmatrix} \frac{1}{4} & \frac{27}{100} & \alpha \\ \frac{27}{100} & -\frac{13}{100} & \alpha \\ \alpha & \alpha & 0 \end{pmatrix}, \begin{pmatrix} -\frac{27}{100} & \frac{21}{100} & 3\alpha \\ \frac{21}{100} & \frac{7}{100} & \alpha \\ 3\alpha & \alpha & 0 \end{pmatrix}, \begin{pmatrix} \frac{7}{50} & -\frac{49}{100} & 3\alpha \\ -\frac{49}{100} & \frac{3}{10} & 0 \\ 3\alpha & 0 & 0 \end{pmatrix} \right)$$

where α is the smallest positive root of the polynomial

$$p(t) = 20828330523 - 3649588559100t^2 + 132250437590000t^4 - 651404153000000t^6 + 748026200000000t^8.$$

Then Y is a non-free matrix extreme point of $\mathcal{D}_A$. We know that p has a positive real zero as $p(0) = 20828330523$ and $p(\frac{1}{8}) = -\frac{208047637414661}{32768}$. So by the intermediate value theorem, there must be a zero between 0 and $\frac{1}{8}$. We go into detail on how this point was computed and proved to be a non-free matrix extreme point of $\mathcal{D}_A$ in Section 5.

4.2. **Exact arithmetic non-free matrix extreme point for $g = 4$.** For $g = 4$, we also have an exact arithmetic example of a non-free matrix extreme point. Let $A = (A_1, A_2, A_3, A_4)$ for

$$A_1 = \operatorname{diag}\left(2, 0, -4, 0, 0, 0, -4, 0, \frac{8}{3}\right), \qquad A_2 = \operatorname{diag}\left(0, 4, -4, 0, 0, 0, 0, -\frac{8}{3}, \frac{8}{3}\right)$$
$$A_3 = \operatorname{diag}\left(0, 0, 0, 4, 0, -\frac{8}{3}, -4, 0, \frac{8}{3}\right), \qquad A_4 = \operatorname{diag}\left(0, 0, 0, 0, \frac{8}{3}, -\frac{8}{3}, 0, -\frac{8}{3}, \frac{8}{3}\right)$$

where $\operatorname{diag}(v)$ is the diagonal matrix whose diagonal is the vector v, be the defining tuple of the free spectrahedron $\mathcal{D}_A$. Then the tuple $X = (X_1, X_2, X_3, X_4)$ for

$$X_1 = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{3}{10} \end{pmatrix} \qquad X_2 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{\frac{3}{5}}}{4} \\ \frac{\sqrt{\frac{3}{5}}}{4} & -\frac{1}{5} \end{pmatrix} \qquad X_3 = \begin{pmatrix} \frac{1521520\sqrt{3}-619599\sqrt{182}}{1019200\sqrt{3}-1197204\sqrt{182}} & 0 \\ 0 & -\frac{1}{4} \end{pmatrix}$$
$$X_4 = \begin{pmatrix} \frac{5\left(1664\sqrt{546}-124455\right)}{3143688} & -\frac{4\left(1820\sqrt{15}+669\sqrt{910}\right)}{392961} \\ -\frac{4\left(1820\sqrt{15}+669\sqrt{910}\right)}{392961} & \frac{11200\sqrt{546}-429603}{3143688} \end{pmatrix}$$

is a non-free matrix extreme point of $\mathcal{D}_A$.

The tuple X was computed by first taking an interior point of $\mathcal{D}_A$ with rational entries, and then perturbing each entry in order to increase the kernel dimension $k_{A,X}$ to a suitable size. We see the result of this method in the structure of X, as each of the X_i's is progressively more complicated.

Of note in this example is that the defining tuple A is a tuple of diagonal matrices, and thus $\mathcal{D}_A$ is a free polytope. It is thanks to this that we can use exact arithmetic to verify that the tuple $X = (X_1, X_2, X_3, X_4)$ is non-free matrix extreme.

5. Exact Arithmetic Point Generation

This section concerns methods for producing provable examples of non-free matrix extreme points X.

We look at two "exact" methods. The first has been effective at producing lots of examples X when $g = 3$, $d = 4$, and $n = 3$. The second though, run extensively for $g = 2$, $d = 3$, and $n = 8$, failed to produce any exact examples, though it produced many examples which are numerically promising. Beyond these parameters, we have not explored either algorithm since our implementations require some intervention, hence are not fast. The example of a non-free matrix extreme point in Section 4 was produced using the first of these exact methods.

5.1. **An algorithm for finding exact arithmetic extreme points.** The algorithm described below generates a boundary point X and then dilates in a precise manner to a point

$$Y = \begin{pmatrix} X & \beta \\ \beta^T & 0 \end{pmatrix}$$

such that $\dim \ker L_A(Y) = 2$. Such a Y is then a good candidate for a non-free matrix extreme point for $g = 3, d = 5$, and $n = 3$, since the Arveson rank-nullity count here is three, and the Matrix extreme rank-nullity count is 2.

Algorithm 5.1. *Let $A \in SM_d(\{-1,0,1\})^g$ such that $\mathcal{D}_A$ is a bounded real free spectrahedron and fix $n \in \mathbb{N}$. Pick a $K \in \{-1,0,1\}^{dn}$ uniformly at random and solve (rational arithmetic) the linear systems*

$$L_A(X)K = 0 \qquad \text{and} \qquad \Lambda_A(\beta^T)K = 0 \tag{5.1}$$

for $X \in SM_n(\mathbb{Q})^g$ and $\beta \in M_{n\times 1}(\mathbb{Q})^g$.

1. *If no solution exists for this K, choose a new $K \in \{-1,0,1\}^{dn}$ uniformly at random and solve (rational arithmetic) the linear systems (5.1).*
2. *If one solution exists for this K, we must then check that $L_A(X)$ is PSD (floating point arithmetic). If $L_A(X)$ is not PSD, we discard this K and choose a new $K \in \{-1,0,1\}^{dn}$ uniformly at random and solve (rational arithmetic) the linear systems (5.1).*
3. *If there are infinitely many solutions, we then pick a tuple X in the solution space such that $L_A(X)$ is PSD.*
4. *Once we have such an X and β, we let*

$$\widehat{Y}(\widehat{\alpha}) = \begin{pmatrix} X & \widehat{\alpha}\beta \\ \widehat{\alpha}\beta^T & 0 \end{pmatrix}$$

   where $\widehat{\alpha} \in \mathbb{R}$. Let α denote the smallest non zero root of the derivative $p_1(\widehat{\alpha}) := \frac{d\chi_{\widehat{\alpha}}(t)}{dt}|_{t=0}$ of the characteristic polynomial $\chi_{\widehat{\alpha}}(t)$ of $L_A(\widehat{Y}(\widehat{\alpha}))$.
5. *Denote $p_2(\widehat{\alpha}) := \frac{d^2\chi_{\widehat{\alpha}}(t)}{d^2t}|_{t=0}$. If $p_2(\widehat{\alpha}) = 0$, then we generate a new X and K and repeat the process. Otherwise return*

$$Y := \widehat{Y}(\alpha).$$

The algorithm uses exact arithmetic so that A, Y, and K have entries which are algebraic numbers.

Theorem 5.2. *If the above algorithm terminates, then the point Y it returns will belong to $\mathcal{D}_A$ and have $k_{A,Y} = 2$. Hence, if $g = 3$, $d = 4$, and $Y \in SM_3(\mathbb{R})^3$, we have that the Arveson Equation (2.1) has more unknowns than constraints, so Y is not an AEP.*

Thus, any point produced with this algorithm at these parameters cannot be an AEP, but is potentially a MEP. The matrix $L_A(Y)$ will have entries that are algebraic numbers, and so the kernel and MEP equations can, in principle, be computed in exact arithmetic.

Proof. Let X, β, K, and α be given by the algorithm above and let

$$Y = \begin{pmatrix} X & \alpha\beta \\ \alpha\beta^T & 0 \end{pmatrix}.$$

We know that

$$L_A(Y) = \Pi^T \begin{pmatrix} L_A(X) & \alpha\Lambda_A(\beta) \\ \alpha\Lambda_A(\beta^T) & I \end{pmatrix} \Pi$$

for some unitary Π (namely the canonical shuffle), so letting $K' = \Pi^T(K^T\ 0)^T$ gives $L_A(Y)K' = 0$ as

$$L_A(Y)K' = \Pi^T \begin{pmatrix} L_A(X) & \alpha\Lambda_A(\beta) \\ \alpha\Lambda_A(\beta^T) & I \end{pmatrix} \Pi K' = \Pi^T \begin{pmatrix} L_A(X)K & 0 \\ \alpha\Lambda_A(\beta^T)K & 0 \end{pmatrix} = 0$$

as $L_A(X)K = \Lambda_A(\beta^T)K = 0$. Thus, the characteristic polynomial $\chi_\alpha(t)$ of $L_A(Y)$ has no constant term. Moreover, by the definition of α, we have $\chi'_\alpha(0) = 0$ and $\chi''_\alpha(0) \neq 0$. Thus $\chi_\alpha(t) = t^2 q_\alpha(t)$ for some polynomial q_α such that $q_\alpha(0) \neq 0$, consequently $L_A(Y)$ has a nullspace of dimension 2.

We now show that $Y \in \mathcal{D}_A$. To do this, we first note that taking $\widehat{\alpha} = 0$, we get that $L_A(\widehat{Y}(0))$ is unitarily equivalent to $L_A(X) \oplus I$ which is clearly PSD. Moreover, since $L_A(\widehat{Y}(\widehat{\alpha}))K' = 0$ for all $\widehat{\alpha}$, we have $\chi_{\widehat{\alpha}}(t) = t(t - \lambda_2(\widehat{\alpha})) \cdots (t - \lambda_{nd+d}(\widehat{\alpha}))$ where the $\lambda_i(\widehat{\alpha})$ are the eigenvalues of $L_A(\widehat{Y}(\widehat{\alpha}))$ that are not identically zero in $\widehat{\alpha}$.

Thus, $p_1(\widehat{\alpha}) = (-1)^{(n+1)d-1}\lambda_2(\widehat{\alpha}) \cdots \lambda_{nd+d}(\widehat{\alpha})$. We note that $p_1(\widehat{\alpha}) = 0$ if and only if $\lambda_i(\widehat{\alpha}) = 0$ for some $i = 2, 3, \ldots, nd + d$ and we pick α to be the smallest positive root of $p_1(\widehat{\alpha})$. Thus, there must be some i such that $\lambda_i(\alpha) = 0$. Moreover, if $\lambda_j(\alpha) < 0$, then, by the intermediate value theorem since $\lambda_j(0) > 0$, there must exist some $0 < \alpha_0 < \alpha$ where $\lambda_j(\alpha_0) = 0$. This implies $p_1(\alpha_0) = 0$, a contradiction to the assumption that α is the smallest positive root of $p_1(\widehat{\alpha})$. Thus for $j \neq i$, $\lambda_j(\alpha) \geq 0$ and hence $Y \in \mathcal{D}_A$.

The Arveson rank-nullity count (2.5) for $g = 3$, $d = 4$, and $n = 3$ is 3. So by Corollary 2.7, if $Y \in \mathcal{D}_A(3)$ then $k_{A,Y} = 2 < 3$ implies that Y is not an AEP. ∎

5.2. **A property of the characteristic polynomial χ_α.**

Lemma 5.3. *If*

$$Y = \begin{pmatrix} X & \alpha\beta \\ \alpha\beta^T & 0 \end{pmatrix}$$

for $X \in SM_n(\mathbb{R})^g$, $\beta \in M_{n\times 1}$, and $\alpha \in \mathbb{R}$, then the characteristic polynomial $\chi_\alpha(t)$ of $L_A(Y)$ has coefficients that are degree dn polynomials in α^2.

Proof. Using the canonical shuffle, we can show that $L_A(Y)$ is unitarily equivalent to

$$Z = \begin{pmatrix} L_A(X) & \alpha\Lambda_A(\beta) \\ \alpha\Lambda_A(\beta^T) & I \end{pmatrix}.$$

The characteristic polynomial of a matrix is invariant under unitary equivalence, so it is sufficient to compute the characteristic polynomial of Z. Note that $Z \in SM_{d(n+1)}(\mathbb{R}[\alpha])$ and thus $\chi_\alpha(t) \in \mathbb{R}[\alpha][t]$.

$$\chi_\alpha(t) = \det \begin{pmatrix} L_A(X) - tI & \alpha\Lambda_A(\beta) \\ \alpha\Lambda_A(\beta^T) & (1-t)I \end{pmatrix}$$

so for $t \neq 1$, we can use the Schur determinant formula to show that

$$\begin{aligned} \det \begin{pmatrix} L_A(X) - tI & \alpha\Lambda_A(\beta) \\ \alpha\Lambda_A(\beta^T) & (1-t)I \end{pmatrix} &= \det((1-t)I) \det([L_A(X) - tI - \frac{\alpha^2}{(1-t)}\Lambda_A(\beta)\Lambda_A(\beta^T)]) \\ &= (1-t)^d \det([L_A(X) - tI - \frac{\alpha^2}{(1-t)}\Lambda_A(\beta)\Lambda_A(\beta^T)]). \end{aligned}$$

Thus $\chi_\alpha(t)$ depends only on α^2 for $t \neq 1$. The matrix in the square brackets above is $dn \times dn$, so this determinant has degree dn in α^2. ∎

5.2.1. *Experiments with Algorithm 5.1.* We ran Algorithm 5.1 one hundred times; every time an X, β, and K were found. Each time the procedure was run, multiple kernels K were generated until an X and β could be found such that $L_A(X)K = 0$ and $\Lambda_A(\beta^T)K = 0$. We have only run this method for a single fixed defining tuple A.

On two randomly chosen occasions, we went through the hour long process of trying to compute the MEP equations exactly. While the equations could in principle be computed in exact arithmetic, the calculation for determining if the MEP equations had a nullspace became too difficult to solve.

We then numerically determined the singular values of the MEP equations, Equation (2.3). In every case, the largest singular value was less than 4 and 80 of the points had smallest singular value on the order of 10^{-4}. Hence, we are confident that those 80 points are MEPs. In order to be AEPs, the points would require a kernel dimension of at least 3. However, in all cases, we computed the eigenvalues of $L_A(X)$ numerically and determined two of the eigenvalues to be zero. This determination was made as the smallest two eigenvalues were on the order of 10^{-15} in all cases,

and the next smallest eigenvalue was on the order of 10^0. Thus, the points could have a kernel of at most dimension 2. Note here that Theorem 5.2 shows that $L_A(Y)$ has kernel dimension 2, so the numerical results match our expectations.

Remark 5.4. *Nowhere in Theorem 5.2 do we use the fact that A, K were chosen from $\{-1, 0, 1\}$. In fact, we may pick A, K with rational entries. The choice to go with $\{-1, 0, 1\}$ entries was made solely to reduce the computational complexity of the experiments.*

5.3. **$g = 2$, the Wild Disc.** For $g = 2$, it is still open whether there exist non-free matrix extreme points. This difficulty is demonstrated by an informal method for the specific example of the wild disc. This method works well to produce floating point non-free matrix extreme candidates but many attempts with exact arithmetic did not yield a provable example. However, we strongly strongly conjecture that non-free matrix extreme points exist for the wild disc.

The $g = 2$ wild disc is the free spectrahedron described by the linear matrix inequality

$$L_A(X, Y) = \begin{pmatrix} 1 & X & Y \\ X & 1 & 0 \\ Y & 0 & 1 \end{pmatrix} \succeq 0,$$

and an easy Schur complement calculation shows that

$$\mathcal{D}_A = \{(X, Y) \mid S(X, Y) := I - X^2 - Y^2 \succeq 0\}.$$

5.3.1. *Finding many non-free matrix extreme points.*

Lemma 5.5. *We have $(X, Y) \in \partial\mathcal{D}_A$ iff $S(X, Y) \succeq 0$ is singular. More precisely,*

$$\ker L_A(X, Y) = \left\{ \begin{pmatrix} v \\ -Xv \\ -Yv \end{pmatrix} \mid v \in \ker S(X, Y) \right\}.$$

Proof. Straightforward. ∎

Based on experimental conclusions described in Section 7.2, we expect that there are pairs (X, Y) of 8×8 matrices that are non-free matrix extreme points of $\mathcal{D}_A$. Indeed, one can attempt to produce many such examples by starting with a random 8×8 matrix $0 \preceq S' \preceq I$ whose rank is $8 - 3 = 5$. Then choose X with $X^2 \preceq I - S'$. All this is easily done with exact arithmetic.

The difficult part is computing exact $Y = (I - X^2 - S')^{1/2}$. In cases where it was easy to compute exact Y, we found that (X, Y) were not non-free matrix extreme. The approaches we used to generate easy to compute Y typically introduced reducibility or some other degeneracy, hence these points were not non-free matrix extreme. On the other hand with nonzero probability (in informal experiments) the pair (X, Y) computed in floating point was a non-free matrix extreme point.

Alternately, given S' as above, to find X, Y one could pick a $2 \cdot 8 \times 8$ isometry $(W_1^T \; W_2^T)^T$ (e.g. by picking first 8 columns of a 16×16 random unitary matrix) and let

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} := \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} (I - S')^{1/2}.$$

Of course, X', Y' won't be self-adjoint. So we correct for this by multiplying W_i with a symmetry U_i commuting with $|W_i(I - S)^{1/2}|$, where $|R| := (R^T R)^{1/2}$. Then

$$X = U_1|W_1(I - S')^{1/2}|, \quad Y = U_2|W_2(I - S')^{1/2}|$$

are self-adjoint with $I - X^2 - Y^2 = S$. Finding appropriate unitaries U_1, U_2 with exact arithmetic is formidable and pursuing this did not produce an exact non-free matrix extreme point. However, as before with nonzero probability (in our experiments) the floating point pair (X, Y) was a non-free matrix extreme point.

6. Numerical Algorithms for Dilation to Extreme Points

Now we turn to numerical experiments; this section gives the underlying algorithms. This paper will use these algorithms in two different ways. One is to generate numerical candidates for non-free matrix extreme points; this will be the main tool behind Section 7. The second goal is a reliable and accurate algorithm for producing a free Caratheodory expansion of a matrix tuple Y inside a free spectrahedron; this is equivalent to dilating Y to an AEP, see Section 1.7. Success rates and statistics on complication of the dilation are in Section 8. We mention that all of the algorithms described in this section have been implemented and are publicly available in NCSE [EEd+21].

As motivation we elaborate on our non-free matrix extreme point objective. While there exist points that are non-free matrix extreme for free spectrahedra in $SM_n(\mathbb{R})^3$ and $SM_n(\mathbb{R})^4$, as demonstrated in Section 4, there is still the question of how frequently these points occur. To answer this question, we turn from methods of generating exact arithmetic extreme points to generating extreme points numerically. Numerical testing is an effective way of generating and testing a large number of points relatively quickly in order to get an idea (without proof) of whether or not example points are rare.

The most obvious method of generating an extreme point of a free spectrahedron $\mathcal{D}_A$ would be to optimize a random linear functional $\ell(X)$ under the constraint $L_A(X) \succeq 0$. However, in practice a surprisingly large majority of points generated by this method are AEPs and have large $k_{A,X}$, as shown in [EFHY21]. Intuitively, and in fact empirically, we expect that points with large $k_{A,X}$ are more likely to be AEPs. This intuition is based on Corollary 2.7. As such, we wish to employ a method of generating extreme points that yields points with small kernels.

6.1. Extreme point generation via dilation.

We begin by describing an algorithm found in [EFHY21], which tries to dilate a given Y of size n_0 in a free spectrahedron $\mathcal{D}_A$ to an AEP. This is done by computing a series of carefully chosen dilations starting with Y.

We find that the success of this algorithm depends heavily on accuracy in computing $\ker L_A(Y^j)$. Nullspace Purification, a method for achieving much improved accuracy, is introduced and tested here in conjunction with Algorithm 6.1. Using Nullspace Purification results in Algorithm 6.1 becoming reliable for broad classes of problems.

6.2. The workhorse extremal dilation algorithm.

Algorithm 6.1 (Extremal Dilation Algorithm [EFHY21]). *Let $A \in SM_d(\mathbb{R})^g$ be such that $\mathcal{D}_A$ is a bounded free spectrahedron and let an initial point $Y^0 \in \mathcal{D}_A(n_0)$ be given. For integers $j = 0, 1, 2, \ldots$ such that Y^j is not an AEP of $\mathcal{D}_A$, define*

$$Y^{j+1} := \begin{pmatrix} Y^j & c_j \widehat{\beta}^j \\ c_j (\widehat{\beta}^j)^T & \widehat{\gamma}^j \end{pmatrix}$$

where $\widehat{\beta}^j$ is a nonzero solution to

$$\ker L_A(Y^j) \subseteq \ker \Lambda_A(\beta^T), \quad \beta \in M_{n\times 1}(\mathbb{R})^g$$

and where c_j and $\widehat{\gamma}^j$ are solutions to the sequence of maximization problems

$$\begin{aligned} c_j := \quad & \underset{c\in\mathbb{R},\gamma\in\mathbb{R}^g}{\text{Maximizer}} \quad c \\ \text{s.t.} \quad & L_A\begin{pmatrix} Y^j & c\widehat{\beta}^j \\ c(\widehat{\beta}^j)^T & \gamma \end{pmatrix} \succeq 0 \\ \text{and} \quad \widehat{\gamma}^j := \quad & \underset{\gamma\in\mathbb{R}^g}{\text{Maximizer}} \quad \ell(\gamma) \\ \text{s.t.} \quad & L_A\begin{pmatrix} Y^j & c_j\widehat{\beta}^j \\ c_j(\widehat{\beta}^j)^T & \gamma \end{pmatrix} \succeq 0. \end{aligned}$$

Here $\ell : \mathbb{R}^g \to \mathbb{R}$ is a random linear functional.

If Y^j is an AEP, we instead terminate the algorithm.

6.3. **Generating non-free matrix extreme points.** In these tests, we do not aim to generate AEPs, and in fact, we are trying to actively avoid generating AEPs, so an algorithm for generating such points may seem like a strange place to look. However, if Y^0 is the initial point, and Y^j is the point generated by Algorithm 6.1, then there are $j-1$ points, $Y^1, Y^2, \ldots, Y^{j-1}$, that we have generated that are "close" to being AEPs but, crucially, are not. As we describe in this section, these sets of points provide many examples of non-free matrix extreme points.

In the experiments detailed in Section 7, we will employ the following algorithm for numerical extreme point generation.

Algorithm 6.2. *Let $A \in SM_d(\mathbb{R})^g$ such that $\mathcal{D}_A$ is a bounded real free spectrahedron. Given an interior point $X \in \mathcal{D}_A(n_0)$, set $Y^0 = \frac{1}{1-\lambda}X$, where λ is the smallest eigenvalue of $L_A(X)$. This guarantees that Y^0 is a boundary point of $\mathcal{D}_A(n_0)$. We then apply Algorithm 6.1 with the difference that the algorithm terminates if some Y^j is either a MEP or an AEP of $\mathcal{D}_A$.*

Theorem 6.3. *Let $\mathcal{D}_A$ be a bounded free spectrahedron and let $X \in \mathcal{D}_A$. Then, with probability 1 Algorithm 6.2 terminates after some finite number k many steps and Y^k is either a MEP or AEP of $\mathcal{D}_A$.*

Proof. Algorithm 6.1 is the subject of [EFHY21, Proposition 2.8]. In this proposition, the authors show that if $Y^0 \in \mathcal{D}_A(n_0)$ and $A \in SM_d(\mathbb{R})^g$, then with probability one, the algorithm terminates in at most $dilDim(Y^0) \leq gn_0$ steps.

Algorithm 6.2 is almost identical to Algorithm 6.1 with the only differences between the two algorithms being the construction of Y^0 from the initial point and the termination of Algorithm 6.2 in the case that some Y^j is a MEP. Once Y^0 is constructed, we apply Algorithm 6.1 and the theory described in paragraph one gives us a bound on the number of steps with probability one. Moreover, if Algorithm 6.2 terminates after k steps, then either Y^k is a MEP or Y^k is an AEP. ∎

Remark 6.4. *While termination in theory occurs with probability 1, in practice this results in a very high but not perfect success rate, see Section 7, 8.*

Given some Y^j as above, Y^{j+1} can be computed by solving two semidefinite programs. The first optimization computes the c_{j+1} and the second computes the $\widehat{\gamma}^{j+1}$. These semidefinite programs can be solved numerically. The introduction of numerical error into the problem requires some consideration. The first main consideration it that a given dilation step can fail in the sense that the dilation is not a maximal 1-dilation. In this case, we discard this failed dilation step and try again with a different β^{j+1}. The β^{j+1} in question is chosen by first computing a basis for the dilation subspace of X with respect to $\mathcal{D}_A$ and then taking a random convex combination of the basis vectors. Thus unless the dilation subspace dimension of X is one, we expect with probability 1 that the newly generated β^{j+1} will not be a scalar multiple of the original.

In order to prevent an infinite loop, we impose a limit on the number of failed dilation attempts on a single point to some maximum number. For the experiments below, this maximum was taken to be ten, meaning a point could fail to dilate ten times before the process was aborted. This number was chosen to be low as the importance of any particular point out of 10,000 trials is relatively low. In cases where it is important to dilate a particular point of interest to an AEP or MEP, it may be appropriate to take a maximum which is much higher.

Another aspect of the algorithm is that there is a $\gamma \in SM_1(\mathbb{R})^g$ that is generated when we compute c_{j+1}, but this γ is "thrown out" and replaced with $\widehat{\gamma}_{j+1}$ in the Y^{j+1}. In our experiments described below, we omit this second step and keep the original γ. Importantly, Theorem 6.3 does not guarantee the termination of this modified algorithm, but in practice it has been shown to be effective.

Given a bounded free spectrahedron $\mathcal{D}_A$, we can use Algorithm 6.2 to generate many extreme points $X \in \mathcal{D}_A(n)$ such that $k_{A,X}$ is relatively small. By producing a large number of extreme points in such a manner and then counting the number of non-free matrix extreme points, we can get a sense for how common such extreme points are.

6.3.1. *What Do We Call Zero?* There is another issue resulting from the introduction of numerical error, namely, what does it mean for a point X with floating point entries to be an extreme point. We can think of such an X as a sum $X = \widehat{X} + X^\delta$ where $\widehat{X}$ is an extreme point of $\mathcal{D}_A$ and X^δ as some small, nonzero error term. Such an X may not even be in $\mathcal{D}_A$, and is likely not an extreme point. The best result we can hope to achieve is for the entries of X^δ to be very small. Thus, instead of aiming to show that X is extreme, we aim to show that X is close to being extreme in the sense that X^δ is small. We will call such points **extreme candidates**.

There are many similar situations where we must make a decision about what to take to be zero and often a decision about which space is the nullspace of a given matrix. We formalize such decisions using the following definition.

Definition 6.5. *Let the matrix $M \in M_{n\times m}(\mathbb{R})$ with singular values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell$, where $\ell = \min(n, m)$. Let $\epsilon_{mag}, \epsilon_{gap} > 0$ be given (ϵ_{mag} will be referred to as the magnitude tolerance and ϵ_{gap} will be referred to as the gap tolerance throughout this paper). A singular value λ_i, for $i = 2, 3, \ldots, \ell$, is said to be the* **first numerical zero** *of M if all of the following are true*

(1) *for all $j < i$, λ_j is not the first numerical zero*
(2) $\lambda_i < \epsilon_{mag}$
(3) $\lambda_i/\lambda_{i-1} < \epsilon_{gap}$.

In other words, a singular value λ_i is the **first numerical zero** *of M if it is the first singular value to be both smaller than the magnitude tolerance and have a sufficiently large gap between it and the previous singular value.*

A singular value λ_j is called a **numerical zero** *if $\lambda_j \leq \lambda_i$, where λ_i is the first numerical zero. Define a function*

$$\Delta(M, \epsilon_{mag}, \epsilon_{gap}) \text{ to be the index of the first numerical zero of } M \tag{6.1}$$

if one exists and False otherwise. (The function Δ is captured by DetermineNull in NCSE).

We now illustrate this by determining if a point X is an extreme candidate of $\mathcal{D}_A$ using a modified version of Theorem 2.6. Suppose we are attempting to show that X is an AEP candidate. As a reminder, a point $X \in \mathcal{D}_A(n)$ is a AEP if and only if the only solution to the linear equation

$$\Lambda_A(\beta^T)K_{A,X} = 0$$

in the unknown $\beta \in M_{n\times 1}(\mathbb{R})^g$ is $\beta = 0$. Thus, the first step in determining if X is an extreme candidate is to compute the nullspace $K_{A,X}$ using some numerical method, see the upcoming subsections for further discussion and for tolerances used for this. The linear map $\beta \mapsto \Lambda_A(\beta^T)K_{A,X}$

has a matrix representation which we will denote $M_{A,X,Arv}$. Thus the equation above has a nontrivial solution if and only if $M_{A,X,Arv}$ has a nontrivial nullspace. In our experiments we say that the nullspace of $M_{A,X,Arv}$ is trivial if and only if $\Delta(M_{A,X,Arv}, 10^{-15}, 10^{-15}) = False$. We chose this very tight tolerance after significant trial and error.

An important choice is what size eigenvalues of $L_A(X)$ to declare numerically 0, that is, determine the nullspace of $L_A(X) \succeq 0$. We approach this using the function Δ. Specify gap and magnitude tolerance, $\epsilon_{gap}, \epsilon_{mag}$, and select

$$\dim \mathcal{K}_{A,X} := dn - \Delta(L_A(X), \epsilon_{mag}, \epsilon_{gap}) + 1$$

where $X \in SM_n(\mathbb{R})^g$ and $A \in SM_d(\mathbb{R})^g$. Once the dimension is selected we can take $\mathcal{K}_{A,X}$ to be the range of a matrix with floating point entries whose columns are orthonormal eigenvectors of $L_A(X)$ corresponding to the $\dim \mathcal{K}_{A,X}$ smallest (i.e., numerically zero) eigenvalues of $L_A(X)$.

6.3.2. *Nullspace Purification.* Improving the numerical accuracy of computing the nullspace $K_{A,X}$ of $L_A(X)$ can significantly improve the accuracy of the Arveson, Euclidean, and matrix extreme equations. One potential avenue for this is to slightly perturb the extreme candidate X to a tuple X^ϵ so that the first numerical zero of $L_A(X^\epsilon)$ is small. The question then arises as to how to compute such a perturbation.

The following algorithm, which we call **Full Nullspace Purification**, is an algorithm that computes such a perturbation by solving a linear program. Linear programs can be solved quickly and to a high degree of accuracy (compared to SDP) which makes this method particularly effective.

Algorithm 6.6 (Full Nullspace Purification)**.** *Given a bounded free spectrahedron $\mathcal{D}_A$, a (numerical) boundary point X of $\mathcal{D}_A$ and tolerances*[2] *$\epsilon_{mag}, \epsilon_{gap}$ and $\epsilon > 0$. Compute*

$$\dim \mathcal{K}_{A,X} := dn - \Delta(L_A(X), \epsilon_{mag}, \epsilon_{gap}) + 1$$

and compute $\mathcal{K}_{A,X}$, the corresponding nullspace of $L_A(X)$. Recall that the nullspace of $L_A(X)$ is nontrivial, since X is a boundary point.

Let $\eta_\epsilon \in \mathbb{R}$, $Y^\epsilon \in SM_n(\mathbb{R})^g$ be the solution to the linear program

$$\begin{aligned} \eta_\epsilon, Y^\epsilon := \quad & \underset{\eta \in \mathbb{R}\, Y \in SM_n(\mathbb{R})^g}{\text{Minimizer}} \quad \eta \\ \text{s.t.} \quad & \max_{i=1,2,\dots,dn} \left| \left[K_{A,X}^T L_A(X+Y) K_{A,X} \right]_{ii} \right| \leq \eta \\ & \|Y\|_{\max} \leq \epsilon \end{aligned} \tag{6.2}$$

where $\|\cdot\|_{\max} : SM_n(\mathbb{R})^g \to \mathbb{R}$ is defined as $\|W\|_{\max} = \max_{i,j,k} |W_{kij}|$ and W_{kij} is the i,j entry of the kth matrix in the tuple W.

Return $X^\epsilon = X + Y^\epsilon$.

Note you may wish to check that $L(X^\epsilon)$ is PSD, to wit that all its eigenvalues are greater than a given tolerance. In algorithms where this is repeated in an inner iteration, checking positivity of the final answer may suffice.

In certain cases, such as when using dilations to compute the free Caratheodory expansion of a point, the extreme point X on which we are applying Nullspace Purification may be of the form

$$X = \begin{pmatrix} X^0 & \beta \\ \beta^T & \gamma \end{pmatrix}$$

where $X^0 \in SM_{n_0}(\mathbb{R})^g$ is of particular importance. In such a case, it is often undesirable to perturb X^0 even slightly, and thus, we employ the following modified algorithm.

[2]In the experiments reported here the tolerances are set to $\epsilon_{mag} = 10^{-7}, \epsilon_{gap} = 10^{-2}$ and $\epsilon = 10^{-7}$.

Algorithm 6.7 (Frozen Nullspace Purification)**.** *Given $\mathcal{D}_A$ be a bounded free spectrahedron and*

$$X = \begin{pmatrix} X^0 & \beta \\ \beta^T & \gamma \end{pmatrix}$$

on the (numerical) boundary of $\mathcal{D}_A$, where $X^0 \in SM_{n_0}(\mathbb{R})^g$, $\beta \in M_{n_0 \times s}(\mathbb{R})^g$, $\gamma \in SM_s(\mathbb{R})^g$, and tolerances $\epsilon_{mag}, \epsilon_{gap}$ and $\epsilon > 0$. Compute

$$\dim \mathcal{K}_{A,X} := dn - \Delta(L_A(X), \epsilon_{mag}, \epsilon_{gap}) + 1$$

and compute $\mathcal{K}_{A,X}$, the corresponding nullspace of $L_A(X)$. Recall that the nullspace of $L_A(X)$ is nontrivial since X is a (numerical) boundary point.

Let $\eta_\epsilon \in \mathbb{R}$, $Y^\epsilon \in SM_n(\mathbb{R})^g$ be the solution to the linear program

$$\begin{aligned} \eta_\epsilon, Y^\epsilon := \quad & \underset{\eta \in \mathbb{R}\, Y \in SM_n(\mathbb{R})^g}{\text{Minimizer}} \quad \eta \\ \text{s.t.} \quad & \max_{i=1,2,\ldots,dn} \left| \left[K_{A,X}^T L_A(X+Y) K_{A,X} \right]_{ii} \right| \leq \eta \\ & \|Y\|_{\max} \leq \epsilon \\ & Y = \begin{pmatrix} 0_{n_0}^g & \tilde{\beta} \\ \tilde{\beta}^T & \tilde{\gamma} \end{pmatrix}. \end{aligned} \tag{6.3}$$

Return $X^\epsilon = X + Y^\epsilon$.

6.3.3. *Observed Algorithm Properties.* While our implementation of Algorithms 6.6 and 6.7 does not guarantee that the nullspace quality improves and that positivity is maintained, we ran extensive experiments that illustrate this happens in practice. In fact, when applying Nullspace Purification to an extreme candidate $X \in \mathcal{D}_A$ we observe that

$$\Delta(L_A(X), 10^{-7}, 10^{-2}) = \Delta(L_A(X^\epsilon), 10^{-11}, 10^{-11}),$$

illustrating that we can use much tighter nullspace tolerances after Nullspace Purification. The above equality represents a significant improvement in nullspace quality, as the magnitude tolerance ϵ_{mag} for extreme candidates generated using SDP alone can rarely be taken below 10^{-8}. That is, for an extreme candidate X that is the output of an SDP, one will frequently have

$$\Delta(L_A(X), 10^{-8}, 10^{-2}) = False.$$

We note that there are many obvious alternatives to Algorithms 6.6 and 6.7. For example, one could minimize all entries of the matrix $K_{A,X}^T L_A(X+Y) K_{A,X}$. It is easier to provide theoretical guarantees for this alternative approach, which is appealing. However, we experimentally observed that our implementation typically notably outperformed the variations that we tried.

6.3.4. *Using Nullspace Purification to improve Algorithm 6.1 and Algorithm 6.2.* In order to increase the numerical accuracy of the extreme point candidates that we generate using Algorithm 6.1 and Algorithm 6.2, after each dilation step we applied Full Nullspace Purification to the dilated point. We are running two types of experiments. In the first, we use Algorithm 6.2 to generate non-free matrix extreme points without much consideration for the actual points generated. In these experiments, we apply Algorithm 6.6, i.e., Full Nullspace Purification, after each dilation step.

In the second set of experiments we perform, the initial point X^0 plays an important role, as we search for a free Caratheodory expansion of that particular point using Algorithm 6.1. Here, we instead use Frozen Nullspace Purification (freezing X^0) as in Algorithm 6.7.

As previously mentioned, it is in principle possible that a Nullspace Purification step could produce $L_A(X^\epsilon)$ that is not numerically PSD. In our experiments, we checked and found that every final dilation point X satisfied $L_A(X) + 10^{-11} I_{dn} \succeq 0$.

The use of Nullspace Purification greatly decreases the rate of failure in our experiments. This is especially true for $g = 2$. It is amusing to note that when running all experiments described in

Section 7 with identically tight tolerances, but without the use of Nullspace Purification, the test took approximately ten times longer to run, and none of the points succeeded in dilating to any form of extreme point. This is not a fair comparison as the accuracy of SDP is often quite low, but the mere fact that it is not a fair comparison shows the power of Nullspace Purification.

The next two sections describe experiments done using these algorithms and also give data describing their success rate and performance.

7. Experimental behavior of non-free matrix extreme points

We randomly generated various free spectrahedra. The goal was to see if "randomly generated" boundary points Y^0 frequently or seldom dilated to non-free matrix extreme points, where we used Algorithm 6.2 as a tool. The data produced is summarized in 3 tables and high level observations and speculations are in Section 9. One consequence of this numerical study was that it guided discovery of the exact $g = 3$ example, Section 4.1. Indeed this suggested using the parameters $d = 4, n = 3$ and the actual spectrahedron we used.

7.1. **Data From Our Experiments: Guide to the Tables.** In our experiments, we consider three different parameters g, d, and n_0 where the defining tuple $A \in SM_d(\mathbb{R})^g$ and the initial point Y^0, discussed in Algorithm 6.2, is in $SM_{n_0}(\mathbb{R})^g$. For every pair g and n_0, we generated 10,000 extreme point candidates using Algorithm 6.2 with Full Nullspace Purification as in Algorithm 6.6. The defining tuples used were randomly generated irreducible tuples $A \in SM_d(\mathbb{R})^g$ where $d = g, g+1, g+2, g+3$. For each defining tuple, 25 points were generated; 100 defining tuples were generated for every value of d totaling 2,500 points for every g, d, n_0. Moreover, the generated spectrahedra were checked to ensure they were bounded.

Throughout Section 7, our experimental data will be presented in tables with the same format as in Table 1. Importantly, the tables describe the properties of the final points generated by the algorithm, not any intermediates. Thus, in the language of Algorithm 6.2, if $Y^0 \in \mathcal{D}_A(n_0)$ is the initial boundary point and $Y^k \in \mathcal{D}_A(n_0+k)$ is the final point, only Y^k and not Y^j for any $0 \leq j < k$ will be represented in the table. For each table, all of the generated points started at the same level n_0, and the n given in column 3 is the level that the points ended at after being dilated.

For any given g, d, and n,

(1) #MnotA, #Euc, and #Arv columns give how many of the points where non-free matrix extreme, EEP, or AEP candidates respectively;
(2) ArvCT, MatCT column gives the Arveson rank-nullity count and Matrix extreme rank-nullity count in that order;
(3) $K_{A,X}$ columns count the number of points with kernel dimension $k_{A,X} = 1, 2, 3, 4, 5$, and > 5. The second number in this column is the list of all the dilation subspace dimensions of the initial points;
(4) #Fail column counts the number of points which failed to dilate to a EEP.

In one case, a point was determined to be a non-free matrix extreme point candidate, but the kernel dimension was uncertain. This case is marked with a $1*$.

A summary of our conclusions from the experiments can be found in Section 9.

7.2. **MEPs of spectrahedra for $g = 2$.** We now narrow our focus to the case where $g = 2$. In this case, we conducted the experiment described above, generating 2,500 extreme point candidates for each g, d, n_0 where $d = 2, 3, 4, 5$. Extreme point candidate generation for $g = 2$ has a higher failure rate ($\sim 0.7\%$ as opposed to the second highest 0.01%) than any other value of g that we tested. This higher rate of failure is the result of a large number of failed dilation tries which results in each dilation step taking significantly longer with some timing out.

A common theme throughout all of our experiments is that we only find examples of non-free matrix extreme point candidates when the Matrix extreme rank-nullity count is strictly less than the

Arveson rank-nullity count. That is, we only find non-free matrix extreme point candidates when the minimum kernel size of $L_A(X)$ necessary to be a MEP is strictly less than the minimum kernel size necessary to be an AEP. We would expect this behavior if the matrix extreme and Arveson extreme equations were randomly generated, but they do have some structure.

We did, however, encounter parameter ranges where the Matrix extreme rank-nullity count is strictly smaller than the Arveson rank-nullity count, but where no non-free matrix extreme point candidates were found. Stronger still, we find parameters, such as $g = 2$, $d = 3$, $n = 5$, where we find EEPs with kernel dimensions sufficiently large to be called MEPs, yet none of these points were determined to be MEP candidates.

$g = 2$, Starting n_0 = 3,4,5														
g	d	n	#Mat not Arv	#Euc	#Arv	ArvCT, MatCT	$K_{A,X}$						#Fail
							1	2	3	4	5	>5	
2	2	3	0	0	0	3,3	0	0	0	0	0	0	0
		4	0	0	0	4,4	0	0	0	0	0	0	0
		5	0	2,500	2,500	5,5	0	0	0	0	2,500;4	0	0
		6	0	0	0	6,6	0	0	1;8	0	2;6	0	3
		7	0	2498	2498	7,6	0	0	0	0	0	2498	0
		8	0	0	0	8,7	0	0	0	0	0	2	2
		9	0	2497	2497	9,8	0	0	0	0	0	2497	0
	3	3	0	0	0	2,2	4;3	0	0	0	0	0	4
		4	0	1773	1746	3,3	9;5	27;3	1746;3	0	0	0	9
		5	0	680	432	4,3	18;7	14;5	248;3,5	432;3	0	0	32
		6	0	2239	2194	4,4	0	26;7	45;5,7	796;3,5	1398;3,5	0	26
		7	0	1511	1464	5,4	0	0	36;7	47;5,7	1464;5,7	0	36
		8	419	1061	626	6,5	0	0	0	16;7	419;5,7	626	0
		9	0	79	74	6,5	0	0	0	0	5;7	74	0
	4	3	0	6	0	2,2	6;2	0	0	0	0	0	0
		4	0	2494	2494	2,2	19;4	956;2	1537;2	0	0	0	19
		5	0	1816	1786	3,3	19;6	30;4	1786;4	0	0	0	19
		6	0	2409	2372	3,3	0	37;6	1998;4,6	374;4	0	0	0
		7	0	498	447	4,3	0	0	51;6	447;6	0	0	0
		8	0	239	239	4,4	0	0	0	119;6	120;6	0	0
		9	0	0	0	5,4	0	0	0	0	0	0	0
	5	3	0	519	0	2,2	519;1	0	0	0	0	0	0
		4	0	2006	1981	2,2	25;3	828;1	1153;1	0	0	0	0
		5	0	2475	2475	2,2	59;5	727;3	1748;3	0	0	0	59
		6	0	1823	1724	3,3	0	99;5	1724;5	0	0	0	0
		7	0	618	618	3,3	0	0	266;5	352;5	0	0	0
		8	0	0	0	4,3	0	0	0	0	0	0	0
		9	0	0	0	4,3	0	0	0	0	0	0	0

TABLE 1

7.3. **MEPs of Spectrahedra for $g = 3$.** We conducted the experiments analogous to those in Section 7.2 for defining tuples $A \in SM_d(\mathbb{R})^3$ and $d = 3, 4, 5, 6$. As in $g = 2$, for parameters where the Matrix extreme rank-nullity count is equal to the Arveson rank-nullity count, we do not find any non-free matrix extreme point candidates. Aside from this, there are some notable distinctions between the $g = 2$ and $g = 3$ cases.

Firstly, we can see that only three points failed to dilate to extreme for $g = 3$ compared to 209 points for $g = 2$. Due to the relative cleanness of the data, we can see that for a given d and n, we usually find points with only one of two different kernel dimensions. The failed points in the $g = 2$ data made it difficult for this phenomenon to be seen.

Secondly, when $g = 3$ there is a much larger array of parameters at which we find non-free matrix extreme point candidates. With the exception of $g = 3$, $d = 4$, $n = 5$ (and higher values of n which no points dilated to), for every g, d, and n where the Matrix extreme rank-nullity count is strictly

less than the Arveson rank-nullity count we find non-free matrix extreme point candidates. The $d = 4$, $n = 5$ exception may be due to experimental design as all the points that started at $n_0 = 2$ and $n_0 = 3$ only needed one dilation step to reach an AEP or a MEP, and all the points that started at $n_0 = 4$ needed at least 3.

$g = 3$, Starting n_0 = 2,3,4														
g	d	n	#Mat not Arv	#Euc	#Arv	ArvCT, MatCT	$K_{A,X}$						#Fail
							1	2	3	4	5	>5	
3	3	2	0	0	0	2,2	0	0	0	0	0	0	0
		3	0	2,500	2,500	3,3	0	0	2,500;3	0	0	0	0
		4	0	1	0	4,4	0	0	1;6	0	0	0	0
		5	0	2499	2499	5,4	0	0	0	0	2499;6	0	0
		6	0	8	0	6,5	0	0	0	0	8;9	0	0
		7	0	2492	2492	7,6	0	0	0	0	0	2492	0
		8	0	0	0	8,6	0	0	0	0	0	0	0
	4	2	0	0	0	2,2	0	0	0	0	0	0	0
		3	261	2,500	2239	3,2	0	261;2	2239;2	0	0	0	0
		4	0	2256	2256	3,3	0	0	2256;5	0	0	0	0
		5	0	187	187	4,3	0	0	0	187;5	0	0	0
		6	299	2516	2217	5,4	0	0	0	299;5,8	2217;5,8	0	0
		7	34	34	0	6,4	0	0	0	0	34;8	0	0
		8	0	7	7	6,5	0	0	0	0	0	7	0
	5	2	0	119	0	2,2	119;1	0	0	0	0	0	0
		3	0	2381	2381	2,2	0	428;1	1953;1	0	0	0	0
		4	398	2,500	2102	3,2	0	397;4	2102;4	0	0	0	0
		5	0	2165	2163	3,3	0	2;7	2163;7	0	0	0	0
		6	64	335	271	4,3	0	0	64;7	271;7	0	0	0
		7	0	0	0	5,4	0	0	0	0	0	0	0
		8	0	0	0	5,4	0	0	0	0	0	0	0
	6	2	0	2,500	2,500	1,1	2,500;0	0	0	0	0	0	0
		3	0	0	0	2,2	0	0	0	0	0	0	0
		4	0	2,500	2,500	2,2	3;6	401;3	2099;3	0	0	0	3
		5	361	2497	2136	3,2	0	361;6	2136;6	0	0	0	0
		6	0	0	0	3,3	0	0	0	0	0	0	0
		7	0	0	0	4,3	0	0	0	0	0	0	0
		8	0	0	0	4,3	0	0	0	0	0	0	0

TABLE 2

7.4. **MEPs of Spectrahedra for $g = 4$.** We conducted the same experiment described in Section 7.2 for defining tuples $A \in SM_d(\mathbb{R})^4$ and $d = 4, 5, 6, 7$.

Similarly to the $g = 3$ case, the only non-free matrix extreme points that we find occur when the Arveson rank-nullity count is strictly larger than the Matrix extreme rank-nullity count. Moreover, at a majority of parameters where we do find the Matrix extreme rank-nullity count to be strictly smaller than the Arveson rank-nullity count, we do find non-free matrix extreme points.

The Matrix extreme rank-nullity count for $g = 4$, $d = 7$ and $n = 2$ is equal to 1 which is unusual. There is a theorem of Helton, Klep and Volčič [HKV18, Section 7] that says that kernel 1 boundary points form an open dense subset of the boundary of irreducible free spectrahedra. The Arveson rank-nullity count is strictly greater than 1, so they cannot be AEPs. This is evidence for the existence of non-free matrix extreme points for this parameter range, and indeed we find many.

7.5. **Dilating MEPs to AEPs.** In our experiments, we produced a modest number of non-free matrix extreme points. Here we check to see if the small perturbations of these points, resulting from Nullspace Purification, dilate to AEPs. By small perturbation, $\widehat{X}$ of X we mean

$$\|X - \widehat{X}\|_{\max} \leq dilDim \cdot 10^{-7}$$

where the norm $\|\cdot\|_{\max}$ is the entry wise maximum as in Algorithm 6.6.

$g = 4$, Starting $n_0 = 2{,}3$													
g	d	n	#Mat not Arv	#Euc	#Arv	ArvCT, MatCT	$K_{A,X}$						#Fail
							1	2	3	4	5	>5	
4	4	2	0	0	0	2,2	0	0	0	0	0	0	0
		3	0	2,500	2,500	3,3	0	0	2,500;4	0	0	0	0
		4	0	0	0	4,4	0	0	0	0	0	0	0
		5	0	2,500	2,500	5,4	0	0	0	0	2,500;8	0	0
		6	0	0	0	6,5	0	0	0	0	0	0	0
		7	0	0	0	7,5	0	0	0	0	0	0	0
	5	2	0	0	0	2,2	0	0	0	0	0	0	0
		3	104	2,500	2396	3,2	0	104;3	2396;3	0	0	0	0
		4	2424	2424	0	4,3	0	0	2424;7	0	0	0	0
		5	0	70	70	4,3	0	0	0	70;7	0	0	0
		6	1	6	5	5,4	0	0	0	1;7	5;7	0	0
		7	0	0	0	6,4	0	0	0	0	0	0	0
	6	2	0	0	0	2,2	0	0	0	0	0	0	0
		3	0	2,500	2,500	2,2	0	179;2	2321;2	0	0	0	0
		4	0	2380	2380	3,3	0	0	2380;6	0	0	0	0
		5	16	120	104	4,3	0	0	16;6	104;6	0	0	0
		6	0	0	0	4,3	0	0	0	0	0	0	0
		7	0	0	0	5,4	0	0	0	0	0	0	0
	7	2	2,500	2,500	0	2,1	2,500;1	0	0	0	0	0	0
		3	0	0	0	2,2	0	0	0	0	0	0	0
		4	188	2,500	2312	3,2	0	188;5	2311;5	0	0	0	0
		5	0	0	0	3,3	0	0	0	0	0	0	0
		6	0	0	0	4,3	0	0	0	0	0	0	0
		7	0	0	0	4,3	0	0	0	0	0	0	0

TABLE 3

7.5.1. $g = 2$. Of the 438 non-free matrix extreme points that were generated in our experiments for $g = 2$, 81 failed to dilate to AEPs. Of the points that dilated to AEPs, all but eight dilated to an AEP in one step. Those eight points dilated to AEPs in two steps. Out of the original 30,000 points generated in our $g = 2$ experiments, a total of 1,445 points failed to dilate to AEPs.

7.5.2. $g = 3$. Of the 1417 non-free matrix extreme points that were generated in our experiments for $g = 3$, 26 failed to dilate to AEPs. Of the points that dilated to Arveson, all but 88 dilated to an AEP in one step. The 88 remaining points dilated in two steps. Out of the original 30,000 points generated in our $g = 3$ experiments, a total of 159 points failed to dilate to AEPs.

7.5.3. $g = 4$. Of the 5,233 non-free matrix extreme extreme points for $g = 4$, all but 332 dilated to AEPs. All but two points that successfully dilated to an AEP dilated in one step. The remaining two dilated to AEPs in two steps. Out of the original 20,000 points generated in our $g = 4$ experiments, a total of 332 points failed to dilate to AEPs.

7.6. **Summary of experimental findings.** Our experiments explored $g = 2, 3, 4$ with d ranging from $d = g$ to $d = g + 3$ with n varying but never exceeding 8. We primarily focused on these parameter ranges both as a way of supporting the process of constructing exact examples and because for larger values of g, the dilation steps take significantly longer to solve.

From the data we can make a number of empirical observations which wild optimism extends to the forthcoming speculation.

Begin by defining the **strict count parameters** g, d, n to be such that the Arveson rank-nullity count is strictly larger than the Matrix extreme rank-nullity count:

$$\left\lceil \frac{gn}{d} \right\rceil > \left\lceil \frac{(n+1)(g+1)}{2d} - \frac{1}{nd} \right\rceil$$

Speculation 7.1. *Consider a generic element A of $SM_d(\mathbb{R})^g$ for which $\mathcal{D}_A$ is bounded. For $g > 2$ and $d > g$ and at each set of strict count parameters, there exist non-free matrix extreme points. In the converse direction, for all g the non-free matrix extreme points can only occur at strict count parameters.*[3] ■

Evidence for Speculation. For $g = 2, 3, 4$ we only find non-free matrix extreme points at strict count parameters. Conversely, for $g = 3$ in seven of the eight strict count parameters where we found points, we found non-free matrix extreme points; six were found in the experiments described in Section 7.3 and one was found in the experiments described in Section 8 where the algorithm used Frozen Nullspace Purification, freezing Y^0 as opposed to the unfrozen Y^0 used in this section. For $g = 4$ the same was true for six of seven strict count parameters. These points were all found in the experiment described in Section 7.4. ■

When g equals d we do not find any non-free matrix extreme points. In this case, we see that for $g = 2, 3, 4$ all of the points generated at $d = g$ that dilated to AEPs ended at the exact same level n. We do not see this phenomenon happening in general for $d > g$. One could speculate that this holds whenever $g = d$; however, we are cautious. For example the relationship between g, d and the boundedness of $\mathcal{D}_A$ changes when g becomes large. Namely, d can be much smaller than g while still having $\mathcal{D}_A$ bounded if g is large. In contrast, if $g = 2$ or 3 and $d < g$, then $\mathcal{D}_A$ is not bounded.

Now we give more detail on the experiment outcomes.

(1) For $g = 2, 3, 4$, we do find values of d and n where there are non-free matrix extreme points.
(2) **Strict count parameters where we do not find non-free matrix extreme points**
  (a) $g = 2$
    (i) $d = 3$: $n = 5, 7, 9$
    (ii) $d = 4$: $n = 7$
  (b) $g = 3$
    (i) $d = 4$: $n = 5$
  (c) $g = 4$
    (i) $d = 5$: $n = 5$ (only 70 points)
(3) **non-free matrix extreme points** are only observed at
  (a) $g = 2$
    (i) $d = 3$: $n = 8$.
  (b) $g = 3$
    (i) $d = 4$: $n = 3, 6, 7$
    (ii) $d = 4$: $n = 8$ (found in Section 8 experiments)
    (iii) $d = 5$: $n = 4, 6$
    (iv) $d = 6$: $n = 5$
  (c) $g = 4$
    (i) $d = 5$: $n = 3, 4, 6$
    (ii) $d = 6$: $n = 5$
    (iii) $d = 7$: $n = 2, 4$
(4) For $g = 2$, individual tries to make a dilation step had a high rate of failure. This causes our experiments to produce fewer MEPs and AEPs per fixed number of start points.

Aside from these main observations, the following are also of interest.

(1) **Starting points.** In each of our experiments we produce a point at level n (over which we have no control except for the starting level). Bounds on n are:
  (a) for $g = 2$ the largest n we found was 9. Here, the starting size was 5;
  (b) for $g = 3$ the largest n we found was 8. Here, the starting size was 4;
  (c) for $g = 4$ the largest n we found was 6. Here, the starting size was 3.

[3] We would expect this behavior if the equations were randomly generated; however, the systems have structure.

(2) **Dilation Dimension.** Given g, d, n. For randomly generated free g, d-spectrahedron and every random boundary point of size n we found that the dilation dimension is $gn - d$. This is what we would expect from a randomly generated linear equation of the same size as the Arveson Equation (2.1). However, the Arveson equations have structure involving several random parts, so they are not truly random.

(3) **Final dilation size** For each $g = 2, 3, 4$, when $g = d$ we observe that all of the Arveson dilations lie at levels 3, 5, or 7. In the case $g = 2$, this behaviour is likely explained by the the fact that all free extreme point of the spin disk lie at level 1 [EHKM18] together Proposition 3.4 and [Eve21, Theorem 3.7] which together allow us to reduce any $g = d = 2$ free spectrahedron to the spin disk.

To our knowledge, there is no canonical $g = d = 3$ or $g = d = 4$ free spectrahedron that all other free spectrahedra with these parameters reduce to. Moreover, numerical evidence suggests that, when $g > 2$, free spectrahedra with $g = d$ do typically have free extreme points at levels above level 1. Therefore, providing an explanation for this behavior is an open question of interest. This behavior is connected to observations made in the upcoming Section 8, where observe that, when $g = d$, our dilation algorithm reduces the dimension of the dilation subspace by a fixed amount on each iteration.

8. Behavior of the Free Caratheodory Expansion Algorithm

In this section we study extremal dilations of a point Y^0 to an AEP. The first major goal is to shed light on the expected level required to dilate an interior Y^0 to an AEP. While [EH19] provides an upper bound, it is not known in which situations the bound is tight. This section provides numerical evidence that Arveson dilations can often be achieved at significantly lower levels than the theoretical bound given in [EH19]. We accomplish this by studying statistics of a parameter μ which intuitively measures the proportion of the theoretical guarantee of [EH19] that is required by our dilation to reach an Arveson extreme point. As seen in the upcoming Tables 4, 5, and 6, and discussed in Section 8.3, the parameter μ is typically much smaller than 1. This illustrates that there are often cases in which the Arveson dilations can be achieved using significantly smaller dilations than needed in the theoretical guarantee.

The second goal of this section, see Section 8.2, regards accuracy of our dilation algorithms using Nullspace Purification and compares them experimentally with and without Freezing. Here experiments concern Nullspace Purification with Freezing, since no Freezing was studied in Section 7.

For interior points Y^0, we shall see that Algorithm 6.1 using Frozen Nullspace Purification, as in Algorithm 6.7, has a failure rate of $< 0.3\%$ for $g = 3, 4$. This compares not badly to the more aggressive Full Nullspace Purification studied in Section 7. This emphasizes that both types of Nullspace Purification considerably improve the performance of our algorithm.

For boundary points Y^0, we will see that Frozen Nullspace Purification has a success rate ($> 84\%$). This contrasts with Full Nullspace Purification which we saw succeeded $> 99\%$ of the time.

While a lower success rate for Frozen Nullspace Purification can be problematic where dilating the *exact original tuple* Y^0, in many applied/numerical settings the tuple Y^0 in question is already a numerical approximation to an underlying "true" tuple of interest. Further perturbing this numerical approximation by a small amount as in Full Nullspace Purification is unlikely to have a significant impact on the reliability of results. In fact, if one believes that the "true" tuple underlying the approximation Y^0 is a boundary point, then purifying the nullspace of $L_A(Y^0)$ can be viewed as a denoising step which increases the numerical reliability of the original approximation. Thus, the high success rate of Full Nullspace Purification can be enjoyed in many applied/numerical settings.

The authors wish to thank Yi Fu for insights on dilations gotten from earlier algorithms.

8.1. **Behavior of the Dilation Algorithm: Guide to the Tables.** In our experiments, we consider three parameters g, d, and n_0 where the defining tuple $A \in SM_d(\mathbb{R})^g$ and the initial point $Y^0 \in \mathcal{D}_A(n_0) \subseteq SM_{n_0}(\mathbb{R})^g$. For each $d = g,\ g+1,\ g+2,\ g+3$, one hundred randomly generated irreducible tuples $A \in SM_d(\mathbb{R})^g$ were used after verifying the boundedness of $\mathcal{D}_A$. For each defining tuple, 25 initial points were randomly generated, totaling 2,500 points for every triple g, d, n_0.

Our data is summarized in 3 tables, one for $g = 2$, for $g = 3$ and for $g = 4$. We also show histograms that help visualize the data presented in the tables. 3D histograms corresponding to the tables are available in the online appendix.

Our tables use the following notation.

(1) Min (resp. Max, Mean) means the minimum (resp. maximum, mean) number of dilation steps, $n - n_0$, observed in our experiment.
(2) We focus on $n - n_0$ as a fraction of the dilation subspace dimension, hence define

$$\mu(n) := \frac{n - n_0}{gn_0} = \frac{n - n_0}{dilDim}.$$

Mean μ (resp. Std. Dev. μ) describes the mean (resp. standard deviation) of μ for each g, d. This is our main measure of the size n of the dilation we obtain.
(3) Fail means a failure to dilate to an AEP or a MEP. Fails are not counted in the statistics (such as the means, standard deviations, etc.).
(4) MnotA means the point dilated to a MEP, but failed to dilate to an AEP. MnotA are not counted in the statistics (such as the means, standard deviations, etc.).

Interior Point Dilation Statistics									
g	d	n_0	$n - n_0$			Mean μ	Std. Dev. μ	#Fail	#MnotA
			Min	Max	Mean				
2	2	3	3	3	3.	0.5	0.	50	0
		4	4	4	4.	0.5	0.	4	0
		5	5	5	5.	0.5	0.	1	0
	3	3	2	6	2.652	0.442	0.131	206	16
		4	2	8	2.82	0.352	0.132	141	68
		5	3	7	3.81	0.381	0.1	161	169
	4	3	1	3	1.658	0.276	0.133	82	0
		4	2	4	2.351	0.294	0.085	89	0
		5	2	5	2.667	0.267	0.08	241	28
	5	3	1	2	1.457	0.243	0.083	60	0
		4	1	3	1.602	0.2	0.097	128	0
		5	2	4	2.189	0.219	0.048	287	70

TABLE 4. $g = 2$.

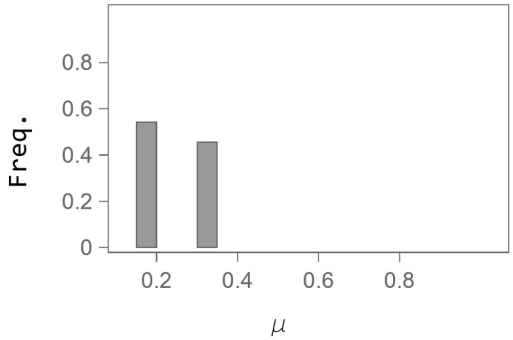

FIGURE 1. To illustrate our histograms, we first highlight the case $g = 2$, $d = 2$, and $n_0 = 3$. The horizontal axis is $0 \leq \mu \leq 1$ and the vertical axis is the probability p with which a value of μ was observed. See Figure 2 for histograms for all cases of d, n_0 when $g = 2$.

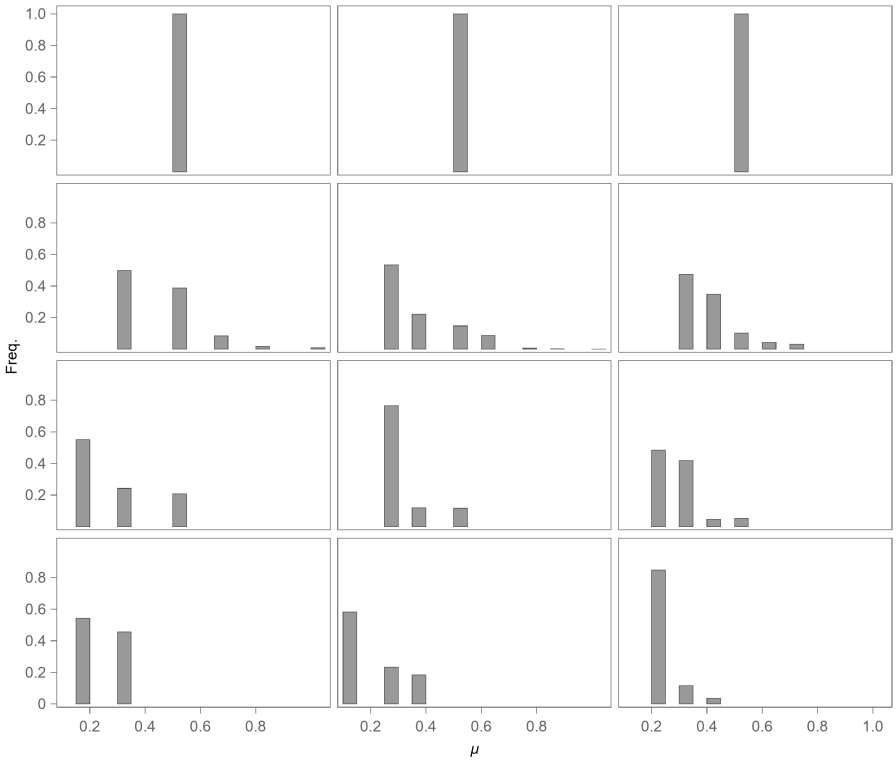

FIGURE 2. These histograms illustrate the data for μ in case of $g = 2$ presented in Table 4. They show the distribution of μ values across various different values of d and n_0. The rows from top to bottom correspond to $d = 2, 3, 4$ respectively and the columns from left to right represent $n_0 = 3, 4, 5, 6$ respectively. Note that for $g = d$, all of the points had a μ value of 0.5.

Interior Point Dilation Statistics									
g	d	n_0	$n-n_0$			Mean μ	Std. Dev. μ	#Fail	#MnotA
			Min	Max	Mean				
3	3	2	2	2	2.	0.333	0.	6	0
		3	3	3	3.	0.333	0.	5	0
		4	4	4	4.	0.333	0.	9	0
	4	2	2	5	2.047	0.341	0.041	2	69
		3	2	9	2.306	0.256	0.083	3	15
		4	3	7	3.241	0.27	0.04	1	36
	5	2	1	3	1.328	0.221	0.101	1	0
		3	2	5	2.066	0.23	0.033	8	0
		4	2	6	2.334	0.195	0.052	7	44
	6	2	1	2	1.26	0.21	0.073	3	0
		3	1	3	1.279	0.142	0.063	2	1
		4	2	4	2.069	0.172	0.024	10	10

TABLE 5. $g = 3$.

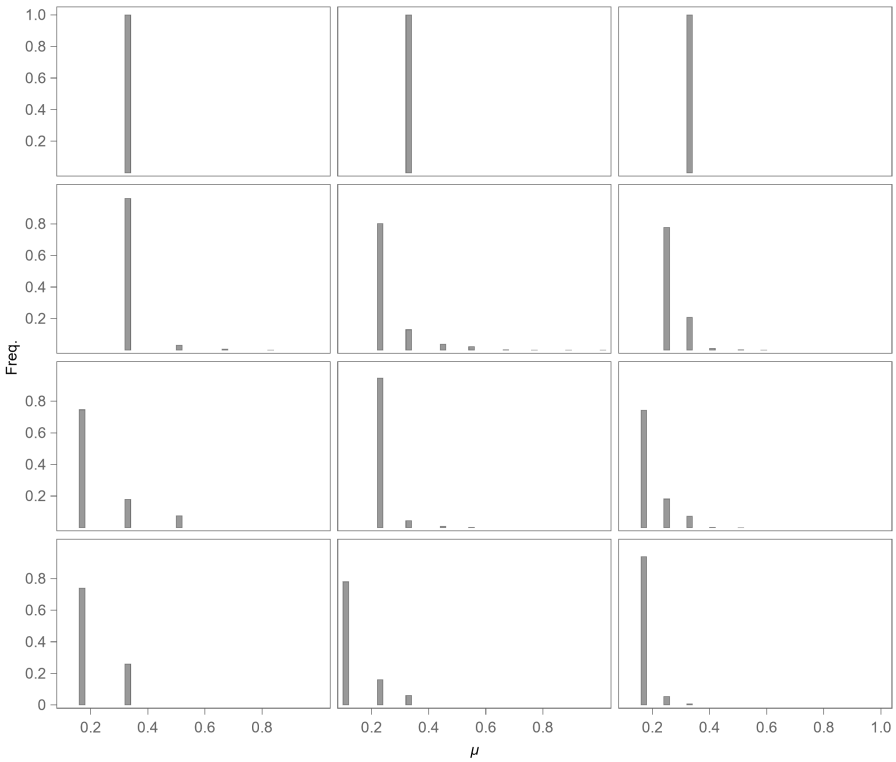

FIGURE 3. These histograms illustrate the data for μ in case of $g = 3$ presented in Table 5. They show the distribution of μ values across various different values of d and n_0. The rows from top to bottom correspond to $d = 3, 4, 5, 6$ respectively and the columns from left to right represent $n_0 = 2, 3, 4$ respectively. Note that for $g = d$, all of the points had a μ value of $1/3$.

Interior Point Dilation Statistics									
g	d	n_0	$n - n_0$			Mean μ	Std. Dev. μ	#Fail	#MnotA
			Min	Max	Mean				
4	4	2	2	2	2.	0.25	0.	8	0
		3	3	3	3.	0.25	0.	1	0
	5	2	2	4	2.088	0.261	0.036	1	17
		3	2	5	2.061	0.172	0.022	0	0
	6	2	1	4	1.11	0.139	0.045	0	0
		3	2	5	2.127	0.177	0.028	2	3
	7	2	1	3	1.137	0.142	0.05	1	0
		3	2	4	2.014	0.168	0.01	2	5

TABLE 6. $g = 4$.

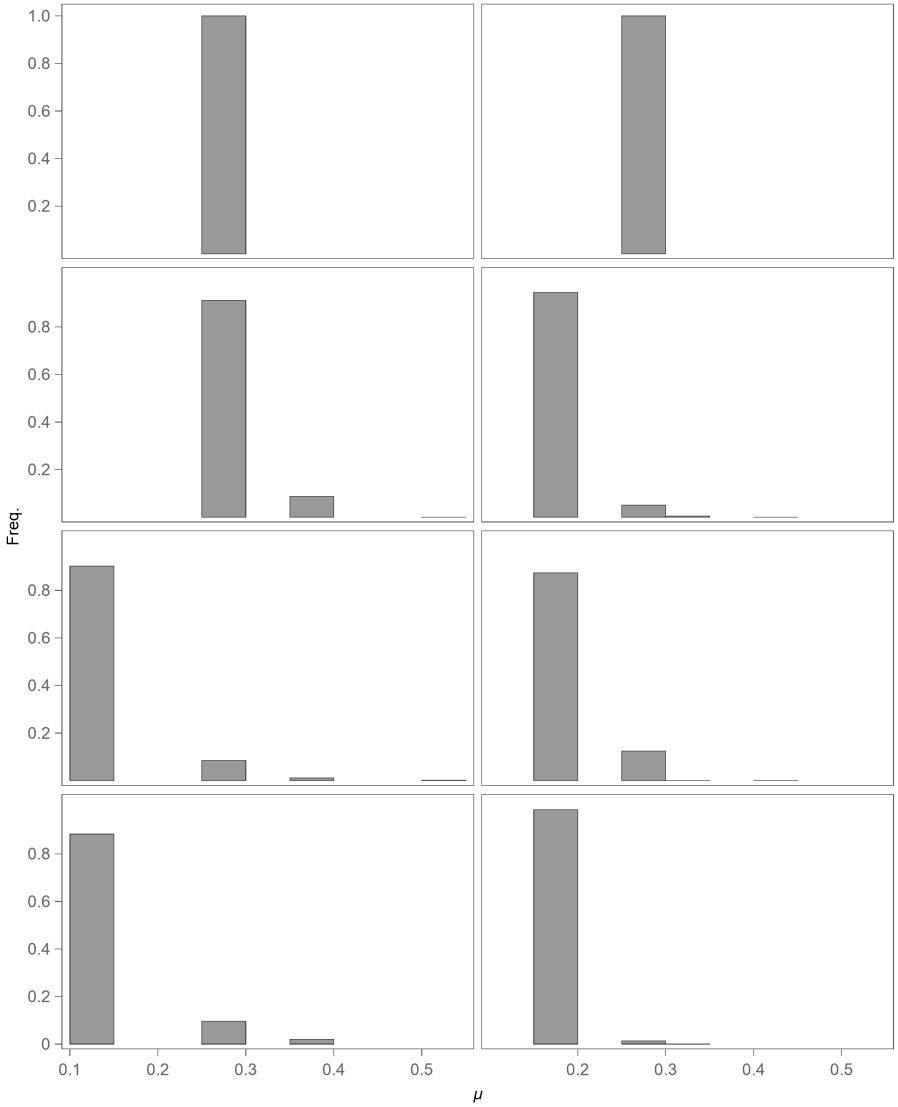

FIGURE 4. These histograms illustrate the data for μ in case of $g = 4$ presented in Table 6. They show the distribution of μ values across various different values of d and n_0. The rows from top to bottom correspond to $d = 4, 5, 6, 7$ respectively and the columns from left to right represent $n_0 = 2, 3$ respectively. Note that for $g = d$, all of the points had a μ value of $1/4$.

8.2. **Summary of Findings with Freezing.**

8.2.1. *Success Rate of Our Algorithm.* Our success rate in dilating to AEPs or MEPs is summarized.

(1) $g = 2$
   (a) Interior Point Failure Rate: $\sim 4.8\%$.
   (b) Boundary Point Failure Rate: $\sim 15.4\%$.
   Here $\sim 8.3\%$ of all initial Y^0 have dilation dimension $= 1$ and
   $\sim 15.7\%$ of those initializations Y^0 which fail.

(2) $g = 3$
   (a) Interior Point Failure Rate: $\sim 0.2\%$
   (b) Boundary Point Failure Rate: $\sim 6.4\%$.
   Here $\sim 8.3\%$ of all initial Y^0 have dilation dimension $= 1$ and
   $\sim 6.8\%$ of those initializations Y^0 which fail.

(3) $g = 4$
   (a) Interior Point Failure Rate: $\sim 0.08\%$
   (b) Boundary Point Failure Rate: $\sim 3.9\%$.

   Here $\sim 8.3\%$ of all initial Y^0 have dilation dimension $= 1$ and
   $\sim 0\%$ of those initializations Y^0 which fail.

8.2.2. *Size of dilations are usually small.*

(1) $Min(n - n_0) \leq Mean(n - n_0) < Min(n - n_0) + \eta_g$
   where $\eta_g < 1$ for all g, d, n_0 and η_g gets smaller as g increases.
(2) As g increases the gap between Max and Min decreases.
(3) The worst case that is theoretically possible μ is 1, but we uniformly find better behavior. For example, when $d - g > 1$ we see the following trend.
   (a) For $g = 2$, the mean of μ is less than 0.3;
   (b) For $g = 3$, the mean of μ is less than 0.23;
   (c) For $g = 4$, the mean of μ is less than 0.18.

8.3. **Size of free extreme dilations and crude insight into to μ.** In a more quantitative vein, a crude estimate for the mean of μ is

$$\mu \sim \mu_{est} := \frac{\lceil (gn_0)/d \rceil}{gn_0}.$$

Note that $\mu = \mu_{est}$ is equivalent to a single dilation step decreasing the dilation subspace dimension by d, so a main use of this formula is as a test for measuring how often this occurs. We compared this (see Table 7) as well as the size of free extreme dilations to the data produced in our experiments. The correspondence of μ_{est} to our experimental data is very rough and typically overestimates the size dilation we get. However, the behavior $\mu = \mu_{est}$ is observed to always hold for $g = d$.

If the dilation subspace equations, Equation (2.1), were generated at random (due to optimizing they are not), one would expect the dilation subspace dimension to be $gn - dk$. Hence, if s is the number of steps taken so far, and we assume that $k = cs$ for some fixed c, we obtain the following equation for the dilation subspace dimension

$$EstDilDim = g(n_0 + s) - d(cs).$$

At the final iteration we have $EstDilDim \leq 0$ which yields $g(n_0 + s) - d(cs) \leq 0$. We can then solve for the smallest integer $\widehat{s}(c)$ such that the inequality holds, giving us

$$\widehat{s}(c) = \left\lceil \frac{gn_0}{cd - g} \right\rceil.$$

Interior Point Dilation Estimate Tests							
g	d	n_0	Error		Mean		
			μ_{est}	%	ΔdilDim/step	ΔdilDim/step past 1	first ΔdilDim
2	2	3	0.	0%	2.	2.	2.
		4	0.	0%	2.	2.	2.
		5	0.	0%	2.	2.	2.
	3	3	0.109	24.7%	2.432	2.015	2.75
		4	-0.023	-6.53%	3.174	3.283	2.917
		5	-0.019	-4.99%	2.773	2.693	2.828
	4	3	-0.057	-20.7%	4.442	3.075	4.2
		4	0.044	15%	3.609	3.153	4.026
		5	-0.033	-12.4%	4.032	3.65	4.304
	5	3	-0.09	-37%	4.628	3.	4.628
		4	-0.05	-25%	6.084	3.9	5.911
		5	0.019	8.68%	4.715	3.59	5.832
3	3	2	0.	0%	3.	3.	3.
		3	0.	0%	3.	3.	3.
		4	0.	0%	3.	3.	3.
	4	2	0.008	2.35%	2.958	1.552	4.345
		3	-0.077	-30.1%	4.137	3.782	4.434
		4	0.02	7.41%	3.768	3.41	4.446
	5	2	-0.112	-50.7%	5.164	3.413	4.985
		3	0.008	3.48%	4.41	2.801	5.997
		4	-0.055	-28.2%	5.411	4.714	6.051
	6	2	0.043	20.5%	5.22	3.	5.22
		3	-0.08	-56.3%	7.923	5.184	7.683
		4	0.005	2.91%	5.869	4.14	7.573
4	4	2	0.	0%	4.	4.	4.
		3	0.	0%	4.	4.	4.
	5	2	0.011	4.21%	3.883	2.037	5.75
		3	-0.078	-45.3%	5.884	5.919	5.812
	6	2	-0.111	-79.9%	7.595	5.603	7.419
		3	0.01	5.65%	5.747	3.987	7.519
	7	2	-0.108	-76.1%	7.506	4.572	7.416
		3	0.001	0.595%	5.972	2.686	9.25

TABLE 7. This table describes how accurate our rough "estimates" are for μ. Error (μ_{est}) is $\mu - \mu_{est}$, so this being negative means that the ultimate dilation is smaller than we would expect based on μ_{est}. The signed percent error is calculated using $(\mu - \mu_{est})/\mu$. ΔdilDim/step is the average of how much the dilation dimension decreases per step.

We would expect $\mu(n) \leq \widehat{s}(1)/(gn_0)$ as the kernel dimension must increase by at least one after every step. This turns out to be true in all of our experiments.

8.3.1. *For each $g = d$ the function $\mu(n)$ is observed to be independent of n.* In our data, we observe that when $g = d$, the dilation dimension decreases by exactly g with each maximal 1-dilation; that is, $\mu(n) = 1/g$. One could reasonably speculate that this trend holds for all spectrahedra for which $g = d$. However, we are cautious in making such a conjecture.

Our caution stems from the fact that, in a sense, there are relatively few different spectrahedra satisfying $g = d$ when g is "small". E.g., Section 3.2 showed that, up to invertible projective maps, there is only one bounded free spectrahedron with $g = d = 2$. It is therefore difficult to say if this behavior is due to the fact that $g = d$ or if it is an artifact which arises when $g = d$ and g is small.

9. Conclusions

In this article we proved the existence of MEPs which are not FEPs for real free spectrahedra. To accomplish this, we produced algorithms that reliably compute, using algebraic arithmetic, non-free matrix extreme points when $g = 3$. In addition when $g = 4$ we have produced "by hand" a non-free matrix extreme point. Furthermore, using theory of projective maps of free spectrahedra, we showed that if $d = 2$, then MEPs and FEPs are the same.

In addition to the above theoretical results, we have "perfected" a numerical algorithm for producing non-free matrix extreme candidates as described in Section 6.1. The reliability of this algorithm relies on a new technique we call Nullspace Purification. This modified algorithm has a much lower rate of failure to the unmodified algorithm used previously [EH19], especially in the $g = 2$ case. The Nullspace Purification algorithm yields extreme point candidates with numerical accuracy on the order of 10^{-13} as opposed to the 10^{-7} accuracy that we find using semidefinite programming alone which allows us to use tighter tolerances when determining if a point is an extreme candidate.

The experiments described in Section 7 yield the following interesting results. Firstly and most importantly, for $g = 2, 3, 4$ we find non-free matrix extreme candidates. Secondly, we never find any non-free matrix extreme candidates when $g = d$. Thirdly, we only find non-free matrix extreme candidates when the Arveson rank-nullity count is strictly greater than the Matrix extreme rank-nullity count. Conversely, for $g = 3, 4$ and $d > g$, in thirteen of sixteen cases, when the Arveson rank-nullity count $\left\lceil \frac{gn}{d} \right\rceil$ is larger than the Matrix extreme rank-nullity count $\left\lceil \frac{gn(n+1)-2}{2dn} \right\rceil$, we find non-free matrix extreme candidates. The three exceptions are potentially explainable through experimental design. We speculate that for $g = 3, 4$ and $d > 4$, if the Arveson rank-nullity count is strictly greater than the Matrix extreme rank-nullity count, then there is a non-free matrix extreme for that value of g, d, n.

Dilating an interior point in a free spectrahedron to an AEP numerically was done using Algorithm 6.1 with Frozen Nullspace Purification and has success rates of

$$95.2\% \text{ for } g = 2, \qquad 99.8\% \text{ for } g = 3, \qquad 99.92\% \text{ for } g = 4.$$

Also, the dilation size was investigated in Section 8. The ratio $\mu(n)$ of the observed dilation step size required to produce an Arveson dilation divided by the theoretical maximum required dilation step satisfies

$$\text{mean}[\mu] < 0.3 \quad \text{for } g = 2, \qquad \text{mean}[\mu] < 0.23 \quad \text{for } g = 3,$$
$$\text{mean}[\mu] < 0.18 \quad \text{for } g = 4.$$

These bounds on the mean of μ show that an Arveson dilation can typically be produced more efficiently than the theoretical maximum required dilation step (which is not itself big) and suggests this effect improves with increasing n.

Data Availability: The data used in this article was randomly generated. The raw data will be made available online https://github.com/NCAlgebra/UserNCNotebooks.

Disclosure statement: The authors report there are no competing interests to declare.

References

[Agl88] J. Agler. An abstract approach to model theory. In *Surveys of some recent results in operator theory, Vol. II*, pages 1–23. Pitman Res. Notes Math. Ser., Longman, Essex, 1988. 3

[AJP20] M. Augat, M.T. Jury, and J.E. Pascoe. Effective noncommutative Nevanlinna-Pick interpolation in the row ball, and applications. *J. Math. Anal. Appl.*, 492(2):124457, 21, 2020. 3

[AM15] J. Agler and J.E. McCarthy. Global holomorphic functions in several noncommuting variables. *Canad. J. Math.*, 67(2):241–285, 2015. 3

[Arv69] W. Arveson. Subalgebras of C^*-algebras. *Acta Math*, 123:141–224, 1969. 3

[Arv08] W. Arveson. The noncommutative Choquet boundary. *J. Amer. Math. Soc.*, 21:1065–1084, 2008. 3

[BGM06] J.A. Ball, G. Groenewald, and T. Malakorn. Bounded real lemma for structured noncommutative multidimensional linear systems and robust control. *Multidimens. Syst. Signal Process.*, 17(2-3):119–150, 2006. 3

[BJN22] A. Bluhm, A. Jenčová, and I. Nechita. Incompatibility in general probabilistic theories, generalized spectrahedra, and tensor norms. *Comm. Math. Phys.*, 393(3):1125–1198, 2022. 3

[BKP16] S. Burgdorf, I. Klep, and J. Povh. *Optimization of Polynomials in Noncommuting Variables*. Springer, 2016. 3

[BN18] A. Bluhm and I. Nechita. Joint measurability of quantum effects and the matrix diamond. *J. Math. Phys.*, 59(11):112202, 27, 2018. 3

[BPT12] G. Blekherman, P.A. Parrilo, and R.R. Thomas. *Semidefinite optimization and convex algebraic geometry*. SIAM, 2012. 1

[DDOSS17] K.R. Davidson, A. Dor-On, M. O. Shalit, and B. Solel. Dilations, inclusions of matrix convex sets, and completely positive maps. *Int. Math. Res. Not.*, 13:4069–4130, 2017. 3

[DK15] K.R. Davidson and M. Kennedy. The Choquet boundary of an operator system. *Duke Math. J.*, 164:2989–3004, 2015. 3

[DlCDN20] G. De las Cuevas, T. Drescher, and T. Netzer. Quantum magic squares: dilations and their limitations. *J. Math. Phys.*, 61(11):111704, 15, 2020. 3

[DLTW08] A.C. Doherty, Y.-C. Liang, B. Toner, and S. Wehner. The quantum moment problem and bounds on entangled multi-prover games. In *Twenty-Third Annual IEEE Conference on Computational Complexity*, pages 199–210. IEEE Computer Soc., Los Alamitos, CA, 2008. 3

[DM05] M. Dritschel and S. McCullough. Boundary representations for families of representations of operator algebras and spaces. *J. Operator Theory*, 53:159–168, 2005. 3

[dOHMP09] M. de Oliveira, J.W. Helton, S. McCullough, and M. Putinar. Engineering systems and free semi-algebraic geometry. In M. Putinar and S.Sullivant, editors, *Emerging applications of algebraic geometry*, pages 17–61. Springer-Verlag, 2009. 3

[dOHMS17] M. de Oliveira, J.W. Helton, B. Miller, and M. Stankus. NCAlgebra 5.04: A noncommutative algebra package for Mathematica, Nov. 2017. 2

[EEd$^+$21] E. Evert, A. Epperly, M. de Oliveira, J. Yin, and J.W. Helton. NCSE 2.3: An NCAlgebra package for optimization over free spectrahedra, Feb. 2021. 2, 18

[EFHY21] E. Evert, Y. Fu, J.W. Helton, and J. Yin. Empirical properties of optima in free semidefinite programs. *To appear in Experimental Mathematics*, 2021. 2, 18, 19

[EH19] E. Evert and J.W. Helton. Arveson extreme points span free spectrahedra. *Math. Ann.*, 375:629–653, 2019. 5, 6, 7, 8, 28, 35

[EHKM18] E. Evert, J.W. Helton, I. Klep, and S. McCullough. Extreme points of matrix convex sets, free spectrahedra and dilation theory. *J. of Geom. Anal.*, 28:1373–1408, 2018. 7, 8, 10, 11, 12, 13, 28

[Eve21] E. Evert. The Arveson boundary of a free quadrilateral is given by a noncommutative variety. *To appear in Operators and Matrices*, 2021. 4, 5, 10, 11, 12, 13, 28

[EW97] E.G. Effros and S. Winkler. Matrix convexity: operator analogues of the bipolar and Hahn-Banach theorems. *J. Funct. Anal.*, 144:117–152, 1997. 1

[Far00] D.R. Farenick. Extremal matrix states on operator systems. *J. London Math. Soc.*, 61:885–892, 2000. 3

[Far04] D.R. Farenick. Pure matrix states on operator systems. *Linear Algebra Appl.*, 393:149–173, 2004. 3

[FHL18] A.H. Fuller, M. Hartz, and M. Lupini. Boundary representations of operator spaces, and compact rectangular matrix convex sets. *J. Operator Theory*, 79:139–172, 2018. 3

[FNT17] T. Fritz, T. Netzer, and A. Thom:. Spectrahedral containment and operator systems with finite-dimensional realization. *SIAM Journal on Applied Algebra and Geometry*, 1(1):556–574, 2017. 3

[GdLL19] S. Gribling, D. de Laat, and M. Laurent. Lower bounds on matrix factorization ranks via noncommutative polynomial optimization. *Found. Comput. Math.*, 19(5):1013–1070, 2019. 3

[Ham79] M. Hamana. Injective envelopes of operator systems. *Publ. Res. Inst. Math. Sci.*, 15:773–785, 1979. 3

[HKM13] J.W. Helton, I. Klep, and S. McCullough. The matricial relaxation of a linear matrix inequality. *Math. Program.*, 138:401–445, 2013. 3, 4

[HKMS19] J.W. Helton, I. Klep, S. McCullough, and M. Schweighofer. Dilations, linear matrix inequalities, the matrix cube problem and beta distributions. *Mem. Amer. Math. Soc.*, 257(1232):vi+106, 2019. 3

[HKV18] J.W. Helton, I. Klep, and J. Volčič. Geometry of free loci and factorization of noncommutative polynomials. *Adv. Math.*, 331:589–626, 2018. 25

[HM12] J.W. Helton and S. McCullough. Every free basic convex semi-algebraic set has an LMI representation. *Ann. of Math.*, 176(2):939–1013, 2012. 3

[HMPV09] J.W. Helton, S. A. McCullough, M. Putinar, and V. Vinnikov. Convex matrix inequalities versus linear matrix inequalities. *IEEE Trans. Automat. Contr.*, 54:952–964, 2009. 3

[JMS21] M.T. Jury, R.T.W. Martin, and E. Shamovich. Non-commutative rational functions in the full Fock space. *Trans. Amer. Math. Soc.*, 374(9):6727–6749, 2021. 3

[Kle14] C. Kleski. Boundary representations and pure completely positive maps. *J. Operator Theory*, 71:45–62, 2014. 3

[Kri19] TL Kriel. An introduction to matrix convex sets and free spectrahedra. *Complex Anal. Oper. Theory*, 13:3251–3335, 2019. 1, 3, 5, 6, 7

[KVV14] D.S. Kaliuzhnyi-Verbovetskyi and V. Vinnikov. *Foundations of free noncommutative function theory*, volume 199 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2014. 3

[MBM21] N.H.A. Mai, A. Bhardwaj, and V. Magron. The constant trace property in noncommutative optimization. In *ISSAC '21—Proceedings of the 2021 International Symposium on Symbolic and Algebraic Computation*, pages 297–304. ACM, New York, [2021] ©2021. 3

[MS98] P.S. Muhly and B. Solel. An algebraic characterization of boundary representations. In *Nonselfadjoint operator algebras, operator theory, and related topics*, Oper. Theory Adv. Appl., pages 189–196. Birkhäuser, Basel, 1998. 3

[MS17] J.A. Mingo and R. Speicher. *Free Probability and Random Matrices*. Springer, 2017. 3

[NT15] T. Netzer and A. Thom. Positivstellensätze for quantum multigraphs. *J. Algebra*, 422:504–519, 2015. 3

[Pas22] B. Passer. Complex free spectrahedra, absolute extreme points, and dilations. *Doc. Math.*, 27:1275–1297, 2022. 3, 6

[PNA10] S. Pironio, M. Navascués, and A. Acín. Convergent relaxations of polynomial optimization problems with noncommuting variables. *SIAM J. Optim.*, 20(5):2157–2180, 2010. 3

[Pop06] G. Popescu. Free holomorphic functions on the unit ball of $B(H)^n$. *J. Funct. Anal.*, 241(1):268–333, 2006. 3

[PR21] V.I. Paulsen and M. Rahaman. Bisynchronous games and factorizable maps. *Ann. Henri Poincaré*, 22(2):593–614, 2021. 3

[RG95] M. Ramana and A.J. Goldman. Some geometric results in semidefinite programming. *J. Global Optim.*, 7:33–50, 1995. 7, 8

[Voi10] D.-V. Voiculescu. Free analysis questions II: the Grassmannian completion and the series expansions at the origin. *J. Reine Angew. Math.*, 645:155–236, 2010. 3

[WM21] J. Wang and V. Magron. Exploiting term sparsity in noncommutative polynomial optimization. *Comput. Optim. Appl.*, 80(2):483–521, 2021. 3

[WW99] C. Webster and S. Winkler. The Krein-Milman theorem in operator convexity. *Trans Amer. Math. Soc.*, 351:307–322, 1999. 3, 5, 6

[Zal17] A. Zalar. Operator Positivstellensätze for noncommutative polynomials positive on matrix convex sets. *J. Math. Anal. Appl.*, 445(1):32–80, 2017. 3

AE: Mathematics Department, UC San Diego, La Jolla, USA
Email address: aepperly@ucsd.edu

EE: Computer Science Department, Northwestern University, Evanston, USA
Email address: eric.evert@northwestern.edu

JWH: Mathematics Department, UC San Diego, La Jolla, USA
Email address: whelton@ucsd.edu

IK: Faculty of Mathematics and Physics, University of Ljubljana & Institute of Mathematics, Physics and Mechanics, Ljubljana Slovenia
Email address: igor.klep@fmf.uni-lj.si