Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Spectrahedral Containment and Operator Systems with finite-dimensional Realization

(revised version: October 2017)

by

Tobias Fritz, Tim Netzer,and Andreas Thom

SPECTRAHEDRAL CONTAINMENT AND OPERATOR SYSTEMS WITH FINITE-DIMENSIONAL REALIZATION

TOBIAS FRITZ, TIM NETZER, AND ANDREAS THOM

ABSTRACT. Containment problems for polytopes and spectrahedra appear in various applications, such as linear and semidefinite programming, combinatorics, convexity and stability analysis of differential equations. This paper explores the theoretical background of a method proposed by Ben-Tal and Nemirovksi [3]. Their method provides a strengthening of the containment problem, that is algorithmically well tractable. To analyze this method, we study abstract operator systems, and investigate when they have a finite-dimensional concrete realization. Our results give some profound insight into their approach. They imply that when testing the inclusion of a fixed polyhedral cone in an arbitrary spectrahedron, the strengthening is tight if and only if the polyhedral cone is a simplex. This is true independent of the representation of the polytope. We also deduce error bounds in the other cases, simplifying and extending recent results by various authors.

1. INTRODUCTION AND PRELIMINARIES

Spectrahedra are the feasible sets of semidefinite programming, and have attracted a lot of attention in recent years, both from an applied and pure perspective (see [4] for an overview). Studying their geometry is a rewarding task for pure mathematicians, but any insight also directly influences the numerous applications, as are optimization, convexity, control theory and others. One of these applications, Lyapunov stability analysis of differential equations, was studied by Ben-Tal and Nemirovski [3]. The problem reduces to checking containment of a box in a spectrahedron, which is a hard problem in general. They came up with a strengthening of this problem, which admits an efficient algorithmic approach. It has been discovered [13] that the method can only be fully understood by adding matricial levels to the spectrahedra, i.e. by examining their free versions. This idea has been further pursued in [7, 12, 17–19].

The results in this paper can be looked at from two sides. On the one hand, we analyze the power of Ben-Tal and Nemirovski's idea, providing a complete description of the cases in which their method provides tight results. In the case of non-tightness we provide error bounds, simplifying and extending upon recent results of several authors. Since tightness is a rather rare phenomenon, such error bounds are of particular interest for applications. Our results show how they directly emerge from geometric properties of the problem, and that they can be computed explicitly. On the other hand, we examine abstract operator systems, and ask when these admit a finite-dimensional concrete realization, i.e. a realization by matrices. This is an interesting and hard problem, which often involves determining the boundary representations of the system (see for example [1, 2, 8]). The connection between the two perspectives becomes clear by observing that free spectrahedra are essentially the same as operator systems with finite-dimensional realizations. We believe that only

the fusion of these two views allows to fully understand the setup of Ben-Tal and Nemirovski's seminal approach, and further exploit the capabilities of spectrahedra in applications.

Our paper is structured as follows. We start with an abstract operator system and characterize when it admits a finite-dimensional realization (Theorem 2.3). We then investigate operator systems constructed from convex cones at scalar level, namely the smallest and the largest operator system of a cone. We show that the largest system admits a finite-dimensional realization if and only if the cone is polyhedral (Theorem 3.2), and the smallest system of a polyhedral cone is finite-dimensional realizable if and only if the cone is a simplex (Theorem 4.7). The smallest system of a non-polyhedral cone can also be finite-dimensional realizable (Example 4.10), but this seems to happen very rarely. Now translated into the initial problem of testing inclusion of spectrahedra, Theorem 4.7 says the following. When checking inclusion of a fixed polytope in an arbitrary spectrahedron, the strengthening first introduced by Ben-Tal and Nemirovski [3] is tight for any spectrahedron if and only if the polytope is a simplex (Corollary 5.3). This is true independently of the representation of the polytope. Further, our setup allows to give an easy proof of the existence of scaling factors for inclusion from [8,19], and prove novel bounds for general spectrahedra (see Section 5).

Let us introduce the basic concepts. Throughout, $\mathcal{V}$ denotes a $\mathbb{C}$-vector space with involution $*$, and $\mathcal{V}_h$ is the $\mathbb{R}$-subspace of Hermitian elements. For any $s \geq 1$, the vector space $\mathbb{M}_s(\mathcal{V}) = \mathcal{V} \otimes_{\mathbb{C}} \mathbb{M}_s(\mathbb{C})$ of $s \times s$-matrices with entries from V comes equipped with the canonical involution defined by $(v_{ij})_{i,j}^* := \left(v_{ji}^*\right)_{i,j}$.

Definition 1.1 (e.g. [22, Chapter 13]). An *abstract operator system* $\mathcal{C}$ on $\mathcal{V}$ consists of a closed and salient convex cone $\mathcal{C}_s \subseteq \mathbb{M}_s(\mathcal{V})_h$ for each $s \geq 1$, such that

- $A \in \mathcal{C}_s, V \in \mathbb{M}_{s,t}(\mathbb{C}) \Rightarrow V^* A V \in \mathcal{C}_t$,
- there is $u \in \mathcal{C}_1 \subseteq \mathcal{V}_h$ such that $u \otimes I_s$ is an order unit (or equivalently interior point) of $\mathcal{C}_s$ for all $s \geq 1$.

Remark 1.2. (a) The topology in which each $\mathcal{C}_s$ is required to be closed is understood to be the finest locally convex topology on $\mathcal{V}$.

(b) We usually consider the order unit $u \in \mathcal{C}_1$ to be part of the structure of an operator system (as opposed to a mere property), which means that maps of operator systems are typically required to preserve it.

(c) $(u \otimes I_s) \in \mathcal{C}_s$ is an order unit for all s if and only if this holds for $s = 1$. To show this, we start with an arbitrary element $A \in \mathbb{M}_s(\mathcal{V})_h$ and decompose it as $A = \sum_{i=1}^{n} v^{(i)} \otimes M_i$ with $v^{(i)} \in \mathcal{V}_h$ and $M_i \in \mathbb{M}_s(\mathbb{C})_h$. Assuming that $u \in \mathcal{C}_1$ is an order unit, choose $\lambda \in \mathbb{R}$ such that $\pm v^{(i)} + \lambda u \in \mathcal{C}_1$ for all i, and write $M_i = P_i - Q_i$ as a difference of two positive semidefinite matrices. Then

$$\sum_i (v^{(i)} + \lambda u) \otimes P_i + (-v^{(i)} + \lambda u) \otimes Q_i = \sum_i v^{(i)} \otimes M_i + \lambda u \otimes \sum_i (P_i + Q_i)$$

is also in $\mathcal{C}_s$. Thus if $\gamma \geq 0$ is large enough to ensure $\gamma I_s - \sum_i (P_i + Q_i) \geqslant 0$, then $A + \gamma\lambda(u \otimes I_s) \in \mathcal{C}_s$. So $u \otimes I_s$ is indeed an order unit for $\mathcal{C}_s$.

By the Choi–Effros Theorem ([6], see also [22, Chapter 13]), for any abstract operator system $\mathcal{C}$ there is a Hilbert space $\mathcal{H}$ and a $*$-linear mapping $\varphi\colon \mathcal{V} \to \mathbb{B}(\mathcal{H})$

with $\varphi(u) = \mathrm{id}_{\mathcal{H}}$, such that for all $s \geq 1$ and $A \in \mathcal{C}_s$,

$$A \in \mathcal{C}_s \iff (\varphi \otimes \mathrm{id})(A) \geqslant 0.$$

On the right-hand side, we use the canonical identification

$$\mathbb{M}_s(\mathbb{B}(\mathcal{H})) = \mathbb{B}(\mathcal{H}) \otimes_{\mathbb{C}} \mathbb{M}_s(\mathbb{C}) = \mathbb{B}(\mathcal{H}^s)$$

to define positivity of the operator. Such a mapping φ is called a *concrete realization* or just *realization* of the operator system $\mathcal{C}$. A realization φ is necessarily injective, since $\mathcal{C}_1$ does not contain a nontrivial subspace.

Definition 1.3. For $r \in \mathbb{N}$, an abstract operator system $\mathcal{C}$ is r-*dimensional realizable* if there is a realization with $\dim \mathcal{H} = r$. It is *finite-dimensional realizable* if it is r-dimensional realizable for some $r \in \mathbb{N}$.

Now assume that $\mathcal{V}$ is finite-dimensional. After a suitable choice of basis, we can assume $\mathcal{V} = \mathbb{C}^d$ with the canonical involution, and thus $\mathcal{V}_h = \mathbb{R}^d$. Then

$$\mathbb{M}_s(\mathcal{V}) = \mathcal{V} \otimes_{\mathbb{C}} \mathbb{M}_s(\mathbb{C}) = \mathbb{M}_s(\mathbb{C})^d, \quad \mathbb{M}_s(\mathcal{V})_h = \mathrm{Her}_s(\mathbb{C})^d,$$

and a realization of $\mathcal{C}$ just consists of self-adjoint operators $T_1, \ldots, T_d \in \mathbb{B}(\mathcal{H})_h$ with $u_1 T_1 + \cdots + u_d T_d = \mathrm{id}_{\mathcal{H}}$ and

$$(A_1, \ldots, A_d) \in \mathcal{C}_s \iff T_1 \otimes A_1 + \cdots + T_d \otimes A_d \geqslant 0.$$

Finite-dimensional realizability then means that the T_i can be taken to be matrices.

Definition 1.4. A *(classical) spectrahedral cone* is a set of the form

$$\left\{ a \in \mathbb{R}^d \mid a_1 M_1 + \cdots + a_d M_d \geqslant 0 \right\},$$

where $M_1, \ldots, M_d \in \mathrm{Her}_r(\mathbb{C})$ are Hermitian matrices, and $\geqslant 0$ again denotes positive semidefiniteness. For any $s \geq 1$, we define

$$\mathcal{S}_s(M_1, \ldots, M_d) := \left\{ (A_1, \ldots, A_d) \in \mathrm{Her}_s(\mathbb{C})^d \mid M_1 \otimes A_1 + \cdots + M_d \otimes A_d \geqslant 0 \right\}.$$

The family of cones $\mathcal{S}(M_1, \ldots, M_d) = (\mathcal{S}_s(M_1, \ldots, M_d))_{s \geq 1}$ is called the *free spectrahedron* defined by $M_1, \ldots, M_d$.

Remark 1.5. In order for a free spectrahedron to be an operator system, the positive cones must be salient and have an order unit. The first is equivalent to the M_i being linearly independent, and the latter happens in particular if there is $u \in \mathbb{R}^d$ with $\sum_i u_i M_i = I_r$, in which case we take this u to be the order unit.

Classical spectrahedra are the feasible sets of semidefinite programming, which allows for efficient numerical algorithms (see for example [25, 26]). They share many properties of polytopes, which form a strict subclass. It is generally hard to decide whether a cone is spectrahedral, and a lot of recent research deals with questions arising in this area (see [4] for an overview). For example, the *inclusion problem* in its basic form asks whether

$$\mathcal{S}_1(M_1, \ldots, M_d) \subseteq \mathcal{S}_1(N_1, \ldots, N_d)$$

holds for given families of matrices M_i and N_j. In Section 5 we will explain how this problem relates to our results. For the moment, just note that a free spectrahedron with the properties of Remark 1.5 is (up to isomorphism) the same as a finite-dimensional realizable operator system.

2. A criterion for finite-dimensional realizations

In this section, we prove a criterion for operator systems to admit a finite-dimensional realization, namely Theorem 2.3 below. Throughout, let $\mathcal{C} = (\mathcal{C}_s)_{s \geq 1}$ be an operator system on $\mathcal{V} = \mathbb{C}^d$ with order unit $u = (u_1, \ldots, u_d) \in \mathcal{C}_1$. Let $\mathcal{C}_s^{\vee}$ denote the dual cone of $\mathcal{C}_s$, i.e. the set of all $*$-linear functionals on $\mathbb{M}_s(\mathbb{C})^d$ that are nonnegative on $\mathcal{C}_s$. We begin by reviewing the separation method of Effros and Winkler.

Lemma 2.1 ([9]). *Let $\varphi \in \mathcal{C}_r^{\vee}$ be such that $\varphi(u \otimes vv^*) > 0$ for all $0 \neq v \in \mathbb{C}^r$. Then there are $M_1, \ldots, M_d \in \mathrm{Her}_r(\mathbb{C})$ with $\sum_i u_i M_i = I_r$, which generate a free spectrahedron containing $\mathcal{C}$, and such that:*

(a) If $A \in \mathrm{Her}_s(\mathbb{C})^d$ is such that $\varphi(V^*AV) = 0$ for some $V \neq 0$, then A is in the boundary of this free spectrahedron.

(b) If $A \in \mathrm{Her}_r(\mathbb{C})^d$ is such that $\varphi(A) < 0$, then A is not in this free spectrahedron.

Proof. Let the $N_1, \ldots, N_d \in \mathrm{Her}_r(\mathbb{C})$ be such that $\varphi(B_1, \ldots, B_d) = \sum_i \mathrm{tr}\left(\overline{N}_i B_i\right)$ for all $B \in \mathbb{M}_r(\mathbb{C})^d$. The positivity assumption guarantees that $\sum_i u_i N_i \geq 0$. Even better, the assumption $\varphi(u \otimes vv^*) > 0$ for all $0 \neq v \in \mathbb{C}^r$ implies that $\hat{N} := \sum_i u_i N_i > 0$, and thus we can put $M_i := \hat{N}^{-1/2} N_i \hat{N}^{-1/2}$ and have $\sum_i u_i M_i = I_r$ by construction.

To show that the resulting free spectrahedron contains $\mathcal{C}$, consider $A \in \mathcal{C}_s$. Then for $x = \sum_{j=1}^r e_j \otimes v_j$ with $v_1, \ldots, v_r \in \mathbb{C}^s$ and $e_1, \ldots, e_r$ the standard basis of $\mathbb{C}^r$, we have

$$(1) \qquad \left\langle x, \left(\sum_i N_i \otimes A_i\right) x \right\rangle = \sum_i \mathrm{tr}\left(\overline{N}_i V^* A_i V\right) = \varphi(V^*AV) \geq 0,$$

where V is the matrix with $v_1, \ldots, v_r$ as its columns. Therefore $\sum_i N_i \otimes A_i \geqslant 0$, which also implies $\sum_i M_i \otimes A_i \geqslant 0$, as was to be shown.

If $\varphi(V^*AV) = 0$ for some $V \neq 0$, then A lies in the boundary of the free spectrahedron, since (1) shows that $\sum_i N_i \otimes A_i$ and hence also $\sum_i M_i \otimes A_i$ is not positive definite, resulting in (a). Part (b) works similarly. $\square$

Definition 2.2. The *essential boundary* of $\mathcal{C}$ is:

$$\partial^{\mathrm{ess}} \mathcal{C}_s := \left\{ A \in \mathcal{C}_s \mid \exists \varphi \in \mathcal{C}_s^{\vee}, \; \varphi(u \otimes vv^*) > 0 \text{ for all } v \in \mathbb{C}^s \setminus \{0\}, \; \varphi(A) = 0 \right\}.$$

So an element is in the essential boundary if and only if its minimal exposed face does not contain an element $u \otimes vv^*$ with $v \neq 0$.

Example 4.2 showcases what the essential boundary of a particular operator system may look like.

Theorem 2.3. *A finite-dimensional operator system $\mathcal{C}$ is r-dimensional realizable if and only if it has the following property: for any $n, s_1, \ldots, s_n \in \mathbb{N}$ and $A^{(i)} \in \partial \mathcal{C}_{s_i}$, there exist $0 \neq V_i \in \mathbb{M}_{s_i, r}(\mathbb{C})$ with*

$$\sum_{i=1}^n V_i^* A^{(i)} V_i \; \in \; \partial^{\mathrm{ess}} \mathcal{C}_r.$$

Proof. First assume that the system is r-dimensional realizable, with defining matrices $M_1, \ldots, M_d \in \mathrm{Her}_r(\mathbb{C})$. For $A^{(i)} = (A_1^{(i)}, \ldots, A_d^{(i)}) \in \partial \mathcal{C}_{s_i}$, there exist vectors $v_1^{(i)}, \ldots, v_r^{(i)} \in \mathbb{C}^{s_i}$ such that

$$x^{(i)} := \sum_{k=1}^{r} e_k \otimes v_k^{(i)} \neq 0,$$

and

$$\left(\sum_{j=1}^{d} M_j \otimes A_j^{(i)} \right) x^{(i)} = 0.$$

Let V_i be the matrix with columns $v_1^{(i)}, \ldots, v_r^{(i)}$. Then $V_i \neq 0$, and some calculation analogous to (1) shows that

$$\mathrm{tr}\left(\sum_{j=1}^{d} \overline{M}_j \sum_{i=1}^{n} V_i^* A_j^{(i)} V_i \right) = \sum_{i=1}^{n} \left\langle x^{(i)}, \left(\sum_{j=1}^{d} M_j \otimes A_j^{(i)} \right) x^{(i)} \right\rangle = 0.$$

This proves that $\sum_i V_i^* A^{(i)} V_i \in \partial^{\mathrm{ess}} \mathcal{C}_r$, since the positive functional $B \mapsto \mathrm{tr}(\sum_j \overline{M}_j B_j)$ is strictly positive on each $u \otimes vv^*$ with $v \neq 0$.

For the converse direction, we use one of the key arguments from [14]. Let $A^{(i)} \in \partial \mathcal{C}_{s_i}$ for $i = 1, \ldots, n$ be elements of the boundary. Then the assumption guarantees that there are $V_i \neq 0$ with $\sum_i V_i^* A^{(i)} V_i \in \partial^{\mathrm{ess}} \mathcal{C}_r$. This means that there is $\varphi \in \mathcal{C}_r^\vee$ with $\varphi(u \otimes vv^*) > 0$ for all $v \in \mathbb{C}^s \setminus \{0\}$ and $\varphi\left(\sum_i V_i^* A^{(i)} V_i\right) = 0$. Since $\varphi \in \mathcal{C}_r^\vee$, this implies that $\varphi\left(V_i^* A^{(i)} V_i\right) = 0$ for each i separately. Hence Lemma 2.1 constructs matrices in $\mathrm{Her}_r(\mathbb{C})$ which generate a free spectrahedron containing $\mathcal{C}$, and such that the $A^{(i)}$ are in its boundary.

The existence of the order unit implies that the defining matrices of such a free spectrahedron are uniformly bounded. Therefore the tuples of matrices that define free spectrahedra containing $\mathcal{C}$ and satisfy $\sum_i u_i M_i = I_r$ form a compact set in $\mathrm{Her}_r(\mathbb{C})^d$. We now choose a sequence of boundary elements $A^{(i)} \in \mathcal{C}_{s_i}$ that are dense in the boundary at all matrix levels, and consider the sequence of free spectrahedra associated to all finite initial subsequences. By compactness, this sequence of free spectrahedra containing $\mathcal{C}$ must have an accumulation point. The free spectrahedron described by such an accumulation point again contains $\mathcal{C}$, and every $A^{(i)}$ is in its boundary. We therefore have an r-dimensional realizable system which has the same boundary as $\mathcal{C}$, and thus coincides with $\mathcal{C}$. $\qquad\square$

We will see in Section 4 how this result can be used to show that certain operator systems are not finite-dimensional realizable.

3. The Largest Operator System of a Cone

In this and the next section, we start with a closed salient cone $C \subseteq \mathbb{R}^d$ with order unit u and consider operator systems $(C_s)_{s \geq 1}$ with $C_1 = C$. It is not hard to see

that there is always a smallest and a largest one, as has also been noticed in [23][1]. We start with the largest system:

$$C_s^{\max} := \left\{ (A_1, \ldots, A_d) \in \operatorname{Her}_s(\mathbb{C})^d \mid \forall v \in \mathbb{C}^s \ (v^* A_1 v, \ldots, v^* A_d v) \in C \right\}.$$

We also write $C^{\max}$ as shorthand for the family $(C_s^{\max})_{s \geq 1}$. This system is largest in the sense that for any operator system $(D_s)_{s \geq 1}$ with $D_1 \subseteq C$, we have $D_s \subseteq C_s^{\max}$ for all s.

The following proposition is a technical ingredient for the main result of this section, Theorem 3.2.

Proposition 3.1. *For $M, N \in \operatorname{Her}_s(\mathbb{C})$, define*

$$\lambda_1 := \min\left\{ \lambda \in \mathbb{R} \ \middle| \ \begin{pmatrix} M + \lambda I & N \\ N & I \end{pmatrix} \geqslant 0 \right\}$$

and

$$\lambda_2 := \min\left\{ \lambda \in \mathbb{R} \mid |w_1|^2 M + 2\operatorname{Re}(w_1 \overline{w}_2) N + (\lambda |w_1|^2 + |w_2|^2) I \geqslant 0 \quad \forall w \in \mathbb{C}^2 \right\}.$$

Then $\lambda_2 \leq \lambda_1$, and if $\lambda_2 = \lambda_1$, then M and N have a common eigenvector.

Proof. It is well-known that $\lambda_1 = \max_{\|v\|=1}(v^* N^2 v - v^* M v)$. Concerning λ_2, it is easy to see that the inequality

$$|w_1|^2 M + 2\operatorname{Re}(w_1 \overline{w}_2) N + (\lambda |w_1|^2 + |w_2|^2) I \geqslant 0 \quad \forall w \in \mathbb{C}^2$$

is equivalent to

$$M + 2rN + (\lambda + r^2) I \geqslant 0 \quad \forall r \in \mathbb{R},$$

and thus to

$$(v^* N v)^2 \leq v^* M v + \lambda \quad \forall \|v\| = 1.$$

Therefore

$$\lambda_2 = \max_{\|v\|=1} [(v^* N v)^2 - v^* M v].$$

We know that $(v^* N v)^2 = (Nv)^* v v^* (Nv) \leqslant (Nv)^* I (Nv) = v^* N^2 v$ for all $\|v\| = 1$, and thus $\lambda_2 \leq \lambda_1$. Whenever $(v^* N v)^2 = v^* N^2 v$, then $Nv \in \ker(I - vv^*)$, so v is an eigenvector of N. Thus if $\lambda_2 = \lambda_1$, then any v that attains λ_2 must also attain λ_1, and therefore be an eigenvector of N.

We finally show that if $\lambda_2 = \max_{\|v\|=1}[(v^* N v)^2 - v^* M v]$ is attained at some eigenvector v of N, then v is also an eigenvector of M. We assume $\|v\| = 1$ and choose an arbitrary w with $\|w\| = 1$ and $w \perp v$. Consider the smooth function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(\epsilon) := \left(\frac{(v + \epsilon w)^*}{\|v + \epsilon w\|} N \frac{(v + \epsilon w)}{\|v + \epsilon w\|} \right)^2 - \frac{(v + \epsilon w)^*}{\|v + \epsilon w\|} M \frac{(v + \epsilon w)}{\|v + \epsilon w\|}$$

and compute

$$f'(0) = -w^* M v - v^* M w,$$

---

[1]Let us emphasize that the largest operator systems in our paper are called *minimal* in [23], while our smallest ones are called the *maximal* ones of [23]. The reason is that the norm induced by a set-theoretically larger system is smaller, and vice versa. So if one is interested in the comparison with operator spaces, then it makes sense to adopt the conventions of [23]. We decided to stick with the set-theoretic notions, hoping that this is less confusing to our readers.

where the derivative of the first term vanishes since v is an eigenvector of N. Since v attains λ_2, there is a maximum of f at $\epsilon = 0$, and therefore $w^*Mv + v^*Mw = 0$. This means $\mathrm{Re}(v^*Mw) = 0$, and by using $-iw$ in place of w also $\mathrm{Im}(v^*Mw) = 0$. Hence $v^*Mw = 0$ for all w with $w \perp v$, which means that the orthogonal complement of v is invariant under M. But then $\mathbb{C}v$ must also be invariant under M, so that v is an eigenvector of M. $\qquad\square$

We can now prove our main result on largest operator systems:

Theorem 3.2. *The operator system $C^{\max}$ admits a finite-dimensional realization if and only if C is polyhedral.*

Proof. One direction is clear: if $C = \left\{ a \in \mathbb{R}^d \mid \ell_1(a) \geq 0, \dots, \ell_r(a) \geq 0 \right\}$, with linear functionals $\ell_i : \mathbb{R}^d \to \mathbb{R}$ such that $\ell_i(u) = 1$ for all i, then for all $s \geq 1$,

$$C_s^{\max} = \left\{ A \in \mathrm{Her}_s(\mathbb{C})^d \mid (\ell_1 \otimes \mathrm{id})(A) \geqslant 0, \dots, (\ell_r \otimes \mathrm{id})(A) \geqslant 0 \right\},$$

and this gives rise to an r-dimensional realization with diagonal matrices.

We now show that the largest system of a non-polyhedral cone does not admit a finite-dimensional realization. First, we argue that we can restrict to the case $d = 3$. Indeed, every non-polyhedral cone C admits a 3-dimensional linear section through 0 and the order unit u, which is not polyhedral either [20, Theorem 4.7], and a possible finite-dimensional realization of $C^{\max}$ would restrict to a finite-dimensional realization of the largest system over this 3-dimensional intersection-cone. So the case $d = 3$ is enough to deal with. Moreover, we can assume that C itself is spectrahedral, since otherwise there is not even a finite-dimensional realization of any system that coincides with C at scalar level.

Now if $C \subseteq \mathbb{R}^3$ is non-polyhedral but spectrahedral, then there is an isomorphism $\varphi \in \mathrm{GL}_3(\mathbb{R})$ such that $C \cap \varphi(C)$ has nonempty interior, but does not have a face of dimension 2. Indeed, the Zariski closure of the boundary of C is an algebraic variety, and hence there must be a smooth point with strict curvature by non-polyhedrality. A reflection φ at a suitable hyperplane close to such a point will then work. Since $C_s^{\max} \cap D_s^{\max} = (C \cap D)_s^{\max}$ holds for any two cones C and D, and the intersection of two systems with finite-dimensional realization has a finite-dimensional realization, we can thus assume that C does not have a face of dimension 2.

Now assume $M_1, M_2, M_3 \in \mathrm{Her}_r(\mathbb{C})$ are defining matrices for $C^{\max}$ of minimal matrix size r. For any $A = (A_1, A_2, A_3) \in \mathrm{Her}_s(\mathbb{C})^3$, we then have

$$\sum_i M_i \otimes A_i \geqslant 0 \;\Leftrightarrow\; A \in C_s^{\max}$$

$$\Leftrightarrow v^*Av \in C = C_1^{\max} \quad \forall v \in \mathbb{C}^s$$

$$\Leftrightarrow \sum_i M_i \cdot v^*A_iv \geqslant 0 \quad \forall v \in \mathbb{C}^s$$

$$\Leftrightarrow \sum_i w^*M_iw \cdot v^*A_iv \geq 0 \quad \forall v \in \mathbb{C}^s, w \in \mathbb{C}^r$$

$$\Leftrightarrow \left\langle \left(\sum_i M_i \otimes A_i \right) x, x \right\rangle \geq 0 \;\forall\; \text{elementary tensors } x \in \mathbb{C}^r \otimes \mathbb{C}^s.$$

Via a suitable change of basis in $\mathbb{C}^3$, we can arrive at $M_3 = I_r$. The above equivalence then entails that the matrix

$$\begin{pmatrix} M_1 + \lambda I_r & M_2 \\ M_2 & I_r \end{pmatrix}$$

is positive if and only if it is positive on all vectors of the form $\begin{pmatrix} w_1 v \\ w_2 v \end{pmatrix}$ for $v \in \mathbb{C}^r$ and $w = (w_1, w_2) \in \mathbb{C}^2$. Using Proposition 3.1, it follows that M_1 and M_2, and trivially also M_3, have a common eigenvector. Thus we can split off a 1×1-block in each M_i. Since the corresponding linear inequality is not needed in the linear inequalities description of C (because there is no face of dimension 2), it is also redundant in the description of the largest system. This contradicts the minimality of r. $\qquad\square$

4. The Smallest Operator System of a Cone

Again let $C \subseteq \mathbb{R}^d$ be a closed salient convex cone with order unit u. Define

$$C_s^{\min} := \left\{ \sum_i c_i \otimes P_i \;\middle|\; c_i \in C, P \in \mathrm{Her}_s(\mathbb{C}), P \geqslant 0 \right\}.$$

Lemma 4.1. $C^{\min}$ *is the smallest operator system with* $C_1^{\min} = C$.

Proof. It is clear that $C^{\min}$ is contained in any operator system extending C.

It remains to check that each $C_s^{\min}$ is closed. By Caratheodory's theorem, the number of elementary tensors required to reach every $A = \sum_i c_i \otimes P_i$ is uniformly bounded. Hence it is enough to show that the set of elementary tensors $\{c \otimes P : c \in C, P \in \mathrm{Her}_s(\mathbb{C})_+\}$ is closed. By choosing any tensor norm, it follows that the elementary tensors of norm 1 are tensor products of elements of norm 1 and therefore form a compact set. $\qquad\square$

Since it will be a crucial ingredient in our main result, we compute the essential boundary of a particular smallest system:

Example 4.2. Consider the cone over the square, i.e. $C = \mathrm{cc}\{v_1, v_2, v_3, v_4\} \subseteq \mathbb{R}^3$, where

(2) $\qquad v_1 = (1, -1, 1), \quad v_2 = (-1, 1, 1), \quad v_3 = (1, 1, 1), \quad v_4 = (-1, -1, 1),$

and $u = (0, 0, 1)$. For $A_1, A_2, A_3, A_4 \geqslant 0$, we have

$$v_1 \otimes A_1 + v_2 \otimes A_2 + v_3 \otimes A_3 + v_4 \otimes A_4 \in \partial^{\mathrm{ess}} C_s^{\min}$$

if and only there is some $U \in \mathrm{GL}_s(\mathbb{C})$ with

(3) $\qquad\qquad \mathrm{im}(UA_1) \perp \mathrm{im}(UA_2) \ \text{and} \ \mathrm{im}(UA_3) \perp \mathrm{im}(UA_4).$

In fact, assume $\varphi \colon \mathrm{Her}_s(\mathbb{C})^3 \to \mathbb{R}$ is nonnegative on $C_s^{\min}$. Then

$$\varphi(X) = \mathrm{tr}(X_1 M_1 + X_2 M_2 + X_3 M_3)$$

for some $M \in \mathrm{Her}_s(\mathbb{C})^3$ with

$$\pm M_1 \pm M_2 + M_3 \geqslant 0$$

for all four sign combinations. Furthermore, $\varphi(u \otimes vv^*) > 0$ for all $0 \neq v \in \mathbb{C}^s$ just means that $M_3 > 0$. So there is some $U \in \mathrm{GL}_s(\mathbb{C})$ with $(U^{-1})^* M_3 U^{-1} = I_s$. Now assume

$$\begin{aligned}
0 &= \varphi(v_1 \otimes A_1 + v_2 \otimes A_2 + v_3 \otimes A_3 + v_4 \otimes A_4) \\
&= \mathrm{tr}(A_1(M_1 - M_2 + M_3) + A_2(-M_1 + M_2 + M_3) \\
&\quad + A_3(M_1 + M_2 + M_3) + A_4(-M_1 - M_2 + M_3)).
\end{aligned}$$

With $S = M_1 + M_2$ and $D = M_1 - M_2$, the above positivity conditions make this equivalent to

$$A_1 \perp (M_3 + D), \quad A_2 \perp (M_3 - D), \quad A_3 \perp (M_3 + S), \quad A_4 \perp (M_3 - S),$$

where we use the standard inner product $\langle X, Y \rangle = \mathrm{tr}(Y^* X)$ on matrices. Thus with $\widetilde{D} := (U^{-1})^* D U^{-1}$,

$$(4) \qquad\qquad U A_1 U^* \perp (I_s + \widetilde{D}), \qquad U A_2 U^* \perp (I_s - \widetilde{D}),$$

and similarly for the other two orthogonality relations involving $\widetilde{S} = (U^{-1})^* S U^{-1}$. Using $-I_s \leq \widetilde{D} \leq I_s$, the spectral decomposition of $\widetilde{D}$, and the fact that eigenvectors to different eigenvalues are orthogonal, we see that $U A_1 U^*$ and $U A_2 U^*$ have orthogonal images, and similarly for $U A_3 U^*$ and $U A_4 U^*$ with S in place of D. This proves (3).

Tracing back this argument, we start with (3), construct $\widetilde{D}$ with spectrum in $[-1, +1]$ such that (4) holds, and similarly for $\widetilde{S}$. This determines M_1, M_2 and M_3 via the above equations, and all desired properties hold by construction.

Before we can prove our main result of this section, we need some more preliminaries.

Definition 4.3. C has a *universal spectrahedral description* of dimension r if there are $M_1, \ldots, M_d \in \mathrm{Her}_r(\mathbb{C})$ with

$$\sum_{i=1}^{d} M_i u_i = I_r, \qquad C = \mathcal{S}_1(M_1, \ldots, M_d),$$

and whenever $N_1, \ldots, N_d \in \mathrm{Her}_t(\mathbb{C})$ with $\sum_i N_i u_i = I_t$, then

$$\mathcal{S}_1(M_1, \ldots, M_d) \subseteq \mathcal{S}_1(N_1, \ldots, N_d) \Rightarrow \forall s \geq 1 : \mathcal{S}_s(M_1, \ldots, M_d) \subseteq \mathcal{S}_s(N_1, \ldots, N_d).$$

This means that the representation detects inclusion of free spectrahedra already at scalar level. This is closely related to realizations of smallest operator systems:

Proposition 4.4. *Let $C \subseteq \mathbb{R}^d$ be a closed salient cone. Then the following are equivalent:*

   (i) The system $C^{\min}$ is finite-dimensional realizable.
   (ii) C admits a universal spectrahedral description.

Proof. (i)$\Rightarrow$(ii): Let $M_1, \ldots, M_d$ realize the system. Whenever

$$C = \mathcal{S}_1(M_1, \ldots, M_d) \subseteq \mathcal{S}_1(N_1, \ldots, N_d),$$

then $\mathcal{S}_s(M_1, \ldots, M_d) = C_s^{\min} \subseteq \mathcal{S}_s(N_1, \ldots, N_d)$ for all s, since the system is the smallest.

(ii)$\Rightarrow$(i): Let $M_1, \ldots, M_d$ be matrices that form a universal spectrahedral description of C. Then $C_s^{\min} \subseteq \mathcal{S}_s(M_1, \ldots, M_d)$ for all $s \geq 1$. Now assume $A \notin C_t^{\min}$ for some $A \in \mathbb{M}_t(\mathbb{C})^d$. Then by choosing a separating positive functional and applying Lemma 2.1, there are $N_1, \ldots, N_d$ with $\sum_i N_i u_i = I_t$ and $C^{\min} \subseteq \mathcal{S}(N_1, \ldots, N_d)$, and such that $A \notin \mathcal{S}_t(N_1, \ldots, N_d)$. From

$$\mathcal{S}_1(M_1, \ldots, M_d) = C = C_1^{\min} \subseteq \mathcal{S}_1(N_1, \ldots, N_d),$$

we obtain $\mathcal{S}(M_1, \ldots, M_d) \subseteq \mathcal{S}(N_1, \ldots, N_d)$ since the description is universal. Thus $A \notin \mathcal{S}_t(M_1, \ldots, M_d)$. We have therefore shown $C_s^{\min} = \mathcal{S}_s(M_1, \ldots, M_d)$ for all $s \geq 1$. $\qquad\square$

Lemma 4.5. *Let $H \subseteq \mathbb{R}^d$ be a subspace that intersects $\mathrm{int}(C)$, and consider the cone $\tilde{C} := C \cap H$. Assume that whenever $\tilde{C} \subseteq \tilde{S} \subseteq H$ for some spectrahedral cone $\tilde{S}$, then the matrix pencil defining $\tilde{S}$ admits an extension to a pencil defining a spectrahedral cone $S \subseteq \mathbb{R}^d$ containing C. If $C^{\min}$ is finite-dimensional realizable, then so is $\tilde{C}^{\min}$.*

Proof. Assume $C^{\min}$ is finite-dimensionally realized by $M_1, \ldots, M_d$. Then the restriction of the pencil spaned by $M_1, \ldots, M_d$ to H yields a universal spectrahedral description of $\tilde{C}$, by the assumed lifting property. In view of Proposition 4.4, $\tilde{C}^{\min}$ is finite-dimensional realizable. $\qquad\square$

Lemma 4.6. *Let $\tilde{C}$ be a face of C. If $C^{\min}$ is finite-dimensional realizable, then so is $\tilde{C}^{\min}$.*

Proof. Since C is in particular spectrahedral, $\tilde{C}$ must be exposed by some $\ell \in C^\vee$ [24]. Then for all matrix levels s and all $A \in \mathbb{M}_s(\mathbb{C})^d$,

$$A \in \tilde{C}_s^{\min} \iff A \in C_s^{\min} \wedge (\ell \otimes \mathrm{id})(A) = 0.$$

Thus every finite-dimensional realization of $C^{\min}$ restricts to a finite-dimensional realization of $\tilde{C}^{\min}$. $\qquad\square$

We are now ready to prove the main result of this section:

Theorem 4.7. *For a salient polyhedral cone $C \subseteq \mathbb{R}^d$, the system $C^{\min}$ is finite-dimensional realizable if and only if C is a simplex. Moreover, $C^{\min} = C^{\max}$ if and only if C is a simplex.*

Proof. One direction is easy. Any simplex cone is isomorphic to the positive orthant $C = \mathbb{R}_{\geq 0}^d$. In this case, one easily checks

$$C_s^{\min} = \left\{ (A_1, \ldots, A_d) \in \mathrm{Her}_s(\mathbb{C})^d \,\middle|\, A_1 \geqslant 0, \ldots, A_d \geqslant 0 \right\} = C_s^{\max}.$$

We prove the other direction in 3 steps.

Step 1: We first deal with the cone over the square $C = \mathrm{cc}\{v_1, v_2, v_3, v_4\} \subseteq \mathbb{R}^3$ as in Example 4.2, and show that its smallest system is not finite-dimensional realizable. This first nontrivial case is already the hardest. We will use Theorem 2.3 together with our characterization of the essential boundary from Example 4.2. Let

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

be the Pauli matrices. For $\alpha \in (0, \pi/2)$, consider the rank one projections

$$A_1 = \tfrac{1}{2}(I_2 - \cos(\alpha)\sigma_z + \sin(\alpha)\sigma_x),$$
$$A_2 = \tfrac{1}{2}(I_2 + \cos(\alpha)\sigma_z - \sin(\alpha)\sigma_x),$$
$$A_3 = \tfrac{1}{2}(I_2 + \cos(\alpha)\sigma_z + \sin(\alpha)\sigma_x),$$
$$A_4 = \tfrac{1}{2}(I_2 - \cos(\alpha)\sigma_z - \sin(\alpha)\sigma_x),$$

where the sign pattern is as in (2), and the associated element

(5) $$A := v_1 \otimes A_1 + v_2 \otimes A_2 + v_3 \otimes A_3 + v_4 \otimes A_4 \in C_2^{\min},$$

still parametrized by $\alpha \in (0, \pi/2)$. For $V \in \mathbb{M}_{r,2}(\mathbb{C})$ with columns w_1, w_2, the property

(6) $$\mathrm{im}(VA_1) \perp \mathrm{im}(VA_2), \quad \mathrm{im}(VA_3) \perp \mathrm{im}(VA_4)$$

is equivalent to $\|w_1\| = \|w_2\|$ and $w_1 \perp w_2$, since A_1 and A_2 are projections onto orthogonal vectors, and likewise for A_3 and A_4. By taking $V = I_2$, or any other unitary, we conclude $A \in \partial^{\mathrm{ess}} C_2^{\min}$ by Example 4.2. Now let $A^{(1)}, \ldots, A^{(r)}$ each be as in (5), but for different angles $0 < \alpha_1 < \ldots < \alpha_r < \frac{\pi}{2}$. If these $A^{(i)}$ admit a compression to $\partial^{\mathrm{ess}} C_r^{\min}$ as in Theorem 2.3, then we obtain $V_i \in \mathbb{M}_{r,2}(\mathbb{C})$ with

$$\mathrm{im}\left(\sum_{i=1}^{r} V_i A_1^{(i)} V_i^*\right) \perp \mathrm{im}\left(\sum_{i=1}^{r} V_i A_2^{(i)} V_i^*\right), \qquad \mathrm{im}\left(\sum_{i=1}^{r} V_i A_3^{(i)} V_i^*\right) \perp \mathrm{im}\left(\sum_{i=1}^{r} V_i A_4^{(i)} V_i^*\right),$$

where now the U of (3) has been absorbed into the V_i. Since each summand is positive, these orthogonality relations require the individual summands to have orthogonal images,

$$\mathrm{im}(V_i A_1^{(i)} V_i^*) \perp \mathrm{im}(V_j A_2^{(j)} V_j^*), \qquad \mathrm{im}(V_i A_3^{(i)} V_i^*) \perp \mathrm{im}(V_j A_4^{(j)} V_j^*),$$

for all $i, j = 1, \ldots, r$. The $A_k^{(i)}$ have rank one, and hence so do the $V_i A_k^{(i)} V_i^*$. An elementary calculation then shows that the $2r$ columns of all the V_i's must be pairwise orthogonal, which is impossible in a space of dimension r. Hence $\sum_i V_i A^{(i)} V_i^*$ cannot be in the essential boundary, and Theorem 2.3 implies that $C^{\min}$ is not r-dimensional realizable. So it is not finite-dimensional realizable. This completes Step 1.

Step 2: We now generalize to those cones D that fit in between the cone over a square and a circumscribed ellipse, as in Figure 1, and will then argue that this actually applies to every salient polyhedral cone in $\mathbb{R}^3$. Again let $C = \mathrm{cc}\{v_1, \ldots, v_4\} \subseteq \mathbb{R}^3$ be as in Step 1. For $\alpha \in (0, \pi/2)$, consider

$$M_1 := \sin(\alpha)\sigma_z, \qquad M_2 := \cos(\alpha)\sigma_x, \qquad M_3 := I_2.$$

Then $C \subseteq C(\alpha) := \{(a, b, c) \in \mathbb{R}^3 \mid aM_1 + bM_2 + cM_3 \geqslant 0\}$. Figure 1 shows sections in the plane defined by $c = 1$ of the cones $C(\alpha)$ for various values of α. Again consider $A \in C_2^{\min}$ as in (5). The functional $(X, Y, Z) \mapsto \mathrm{tr}(XM_1 + YM_2 + ZM_3)$ even shows that $A \in \partial^{\mathrm{ess}} D_2^{\min}$ for any convex cone D with $C \subseteq D \subseteq C(\alpha)$. Now assume that the inclusions
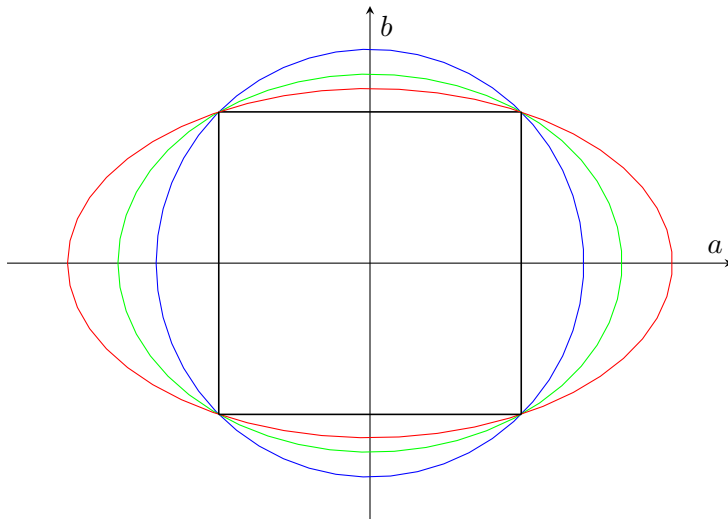
$$C \subseteq D \subseteq C(\alpha)$$

FIGURE 1. A section of the cone C (square) together with some sections of various $C(\alpha)$.

hold for *infinitely* many values of $\alpha \in (0, \pi/2)$. Then families of $A^{(i)}$ as above (with different values for α) are also in $\partial D_2^{\min}$, but cannot be compressed into $\partial^{\mathrm{ess}} D_r^{\min}$ as in Theorem 2.3, since this would then also work for $\partial^{\mathrm{ess}} C_r^{\min}$. Hence the operator system $D^{\min}$ is not finite-dimensional realizable.

To complete Step 2, we show that this applies to *every* salient polyhedral cone $D \subseteq \mathbb{R}^3$. Any quadrilateral in the plane can be transformed by a projective transformation to the square. So in a given planar polytope which is not a simplex, choose vertices u_1, u_2, w_1, w_2 that form a quadrilateral, such that both pairs u_1, u_2 and w_1, w_2 are adjacent vertices. Then transform them to the square, and choose $\alpha' > 0$ such that the transformed polytope is contained in $C(\alpha)$ for all $0 < \alpha < \alpha'$. This is possible, since the gradient to $\det(aM_1 + bM_2 + M_3)$ at $(a, b) = (1, 1)$ tends to $(1, 0)$ for $\alpha \to 0$, and similarly at the other three corners of the square. This shows that any non-simplex polyhedral cone in $\mathbb{R}^3$ is isomorphic to a cone D with $C \subseteq D \subseteq C(\alpha)$ for infinitely many values of $\alpha \in (0, \pi/2)$. Its smallest system is thus not finite-dimensional realizable.

Step 3: We prove the statement in arbitrary dimension $d \geq 4$ by induction on d. If C is not a simplex, then either it has a facet that is not a simplex, or a vertex figure that is not a simplex [27, p. 67]. In the first case we apply the contrapositive of Lemma 4.6, while in the second case we apply Lemma 4.5 to a hyperplane defining the vertex figure. The extension required by Lemma 4.5 is possible by taking the conical hull of $\tilde{S}$ from the vertex (ray). In both cases we reduce to dimension $d - 1$.

Finally, the statement about $C^{\min} = C^{\max}$ follows from the previous results. $\square$

Remark 4.8. (i) The argument in Step 2 of the previous proof shows that the smallest system of many non-polyhedral cones is not finite-dimensional realizable either. Any cone in $\mathbb{R}^3$ having a compact section that contains the square and is contained in two different $C(\alpha)$ is an example.

(ii) The results from [10] provide further evidence that finitely generated operator systems are hardly ever finite-dimensional realizable.

Remark 4.9. Every polyhedral cone can be regarded either as the set of positive linear combinations of its finitely many extreme points, or as the set of all points satisfying its finitely many facet inequalities. $C^{\min}$ extends the first picture to matrix levels, since we take matrix positive combinations of points from C. On the other hand, $C^{\max}$ generalizes the second picture, since it is defined by the inequalities of C. Except for simplices, these two extensions are thus different at matrix level.

Example 4.10. There are non-polyhedral cones with a finite-dimensional realizable smallest operator system. One example is the circular cone
$$C = \big\{(a,b,c) \in \mathbb{R}^3 \mid c \geq 0,\ a^2 + b^2 \leq c^2\big\}.$$
It is proven in [12, Corollary 14.15] and [16, Theorem 5.4.10] (which relies mostly on [5, Theorem 7]), that the following linear matrix pencil defines the smallest system:
$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes x + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes y + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes z.$$
It is tempting to conjecture that the following pencil $\mathcal{L}$ defines the smallest system of the analogous cone over the three-dimensional Euclidean ball in $\mathbb{R}^4$:
$$\mathcal{L}(x,y,w,z) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes x + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes y + \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \otimes w + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes z,$$
which indeed coincides with that cone at the first matrix level. However, this is not true. It is well-known that there are hermitian 4×4-matrices of the block form
$$X = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix}$$
that are positive semidefinite, but cannot be written as $\sum_i P_i \otimes Q_i$ with positive semidefinite matrices P_i, Q_i, where all P_i are of size 2; such matrices are called *entangled* in the language of quantum physics. An easy example is the rank one projection
$$X = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{array} \right).$$
Now $0 \leqslant 2X = \mathcal{L}\big(A - C, B + B^*, \frac{1}{i}(B - B^*), A + C\big)$. So if the inequality $\mathcal{L} \geqslant 0$ defined the smallest operator system of the cone C over the three-dimensional Euclidean ball, then for every X there would be vectors $v_i \in C$ and positive semidefinite matrices Q_i such that
$$\left(A - C, B + B^*, \frac{1}{i}(B - B^*), A + C\right) = \sum_i v_i \otimes Q_i.$$
But then
$$2X = \sum_i \underbrace{\mathcal{L}(v_i)}_{\geqslant 0} \otimes Q_i,$$
which contradicts the possibility that X may be entangled.

5. Inclusion of Spectrahedra

We explain how our results relate to inclusion testing of spectrahedra. The inclusion testing problem is the following:

Problem 5.1. Given $M_1, \ldots, M_d \in \mathrm{Her}_r(\mathbb{C})$ and $N_1, \ldots, N_d \in \mathrm{Her}_t(\mathbb{C})$ with

$$\sum_i u_i M_i = I_r, \qquad \sum_i u_i N_i = I_t,$$

then is it true that

$$\mathcal{S}_1(M_1, \ldots, M_d) \subseteq \mathcal{S}_1(N_1, \ldots, N_d)$$

holds in $\mathbb{R}^d$?

Already if $\mathcal{S}_1(M_1, \ldots, M_d)$ is the cone over a d-dimensional cube, this question arises in interesting applications [3]. In general, it is a hard algorithmic problem (see [19] for an overview and new results). The following strengthening was introduced in [13, Section 4.1], generalizing the strengthening of [3, Eq. (7)] for the matrix cube problem:

Problem 5.2 ([13]). Given $M_1, \ldots, M_d \in \mathrm{Her}_r(\mathbb{C})$ and $N_1, \ldots, N_d \in \mathrm{Her}_t(\mathbb{C})$ with

$$\sum_i u_i M_i = I_r, \qquad \sum_i u_i N_i = I_t,$$

do there exist $V_j \in \mathbb{M}_{r,t}(\mathbb{C})$ such that

$$\sum_j V_j^* M_i V_j = N_i$$

for all i?

A positive answer to an instance of Problem 5.2 implies a positive answer to the corresponding instance of Problem 5.1. Furthermore, Problem 5.2 can be formulated as a semidefinite feasibility problem, and is thus algorithmically tractable. However, a positive answer to Problem 5.1 does not necessarily imply a positive answer to Problem 5.2. The main result of [13] says that Problem 5.2 is equivalent to $\mathcal{S}_s(M_1, \ldots, M_d) \subseteq \mathcal{S}_s(N_1, \ldots, N_d)$ for *all* $s \geq 1$, i.e. to inclusion of the free spectrahedra. This result mostly relies on Choi's characterization of completely positive maps between matrix algebras [5]. Since the inclusion $\mathcal{S}_1(M_1, \ldots, M_d) \subseteq \mathcal{S}_1(N_1, \ldots, N_d)$ does not imply the higher inclusions $\mathcal{S}_s(M_1, \ldots, M_d) \subseteq \mathcal{S}_s(N_1, \ldots, N_d)$ in general, Problem 5.2 is a proper strengthening of Problem 5.1. There exist quantitative measures for tightness of this strengthening [3, 7, 12, 18], which we will explain in more detail below. The first reformulation of our previous result is the following:

Corollary 5.3. *Assume* $C = \mathcal{S}_1(M_1, \ldots, M_d) \subseteq \mathbb{R}^d$ *is a salient polyhedral cone. Then a positive answer to Problem 5.1 implies a positive answer to Problem 5.2 for all choices of* $N_1, \ldots, N_d \in \mathrm{Her}_t(\mathbb{C})$ *if and only if* C *is a simplex.*

Of course, even when C is not a simplex, the strengthening may still give the correct answer for a *particular* choice of $N_1, \ldots, N_d$.

Proof. Problems 5.1 and 5.2 are equivalent for all $N_1, \ldots, N_d$ if and only if the M_i provide a universal spectrahedral description of C. So the result follows from Proposition 4.4 and Theorem 4.7. $\qquad\square$

Remark 5.4. (i) Although the tightness of the strengthening for simplex cones is easy to prove, it seems like it has not been observed in the literature so far. We will use it below to easily derive error bounds for the non-tight case.

(ii) Corollary 5.3 holds for *any description* of C by matrices M_i. So far, only fixed descriptions have been used to deduce error bounds and non-tightness results in [3, 7, 12, 18]. It was not clear a priori whether choosing a better spectrahedral description of the cones could result in tightness. We now know that this is impossible for non-simplex polyhedral cones.

(iii) Finally, we are not aware of any implications of our results on the tightness of the higher levels of the semidefinite hierarchy of [18].

Using our approach, parts of [18, Theorem 4.8] become easy to prove. The result implies that for inclusion of spectrahedra in polyhedra, the strengthening is always tight.

Proposition 5.5. *If $N_1, \ldots, N_d$ commute and $C \subseteq \mathcal{S}_1(N_1, \ldots, N_d)$, then*
$$C^{\max} \subseteq \mathcal{S}(N_1, \ldots, N_d).$$

Proof. $P = \mathcal{S}_1(N_1, \ldots, N_d)$ is polyhedral, and thus $P^{\max} = \mathcal{S}(N_1, \ldots, N_d)$ by the easy direction in the proof of Theorem 3.2. The claim now follows from $C^{\max} \subseteq P^{\max}$. $\square$

As explained above, a positive answer to the strengthened Problem 5.2 implies a positive answer to the original Problem 5.1. In the situation of Corollary 5.3 or Proposition 5.5, also a negative answer to Problem 5.2 implies the same for Problem 5.1, although such an inference is not valid in general. One way to approach this issue is to use an entire hierarchy of semidefinite programs that converge to Problem 5.1 [18]. Another one is to modify the formulation of Problem 5.2 so as to make the implication work; concretely, we can replace the M_i in Problem 5.2 by a "scaled down" version which is small enough for the implication of negative answers to be valid.

For a salient convex cone C, the definition of scaling is as follows. Choose an arbitrary hyperplane H that intersects C only at the origin. For a scaling factor $\nu > 0$, the scaled cone $\nu{\uparrow}C$ is constructed by taking the intersection of C with the affine hyperplane $u + H$, scaling this intersection by the factor ν from the point u, and taking the conical hull with the origin again, resulting in $\nu{\uparrow}C$. In general, $\nu{\uparrow}C$ depends on the choices of H and u.

While $C^{\min} \subseteq C^{\max}$ holds trivially, we can now ask by how much $C^{\min}$ and $C^{\max}$ differ, namely by investigating how small the scaling factor $\nu > 0$ needs to be in order for the reverse inclusion to hold,

(7) $$(\nu{\uparrow}C)^{\max} \subseteq C^{\min}.$$

We now derive some results on this and then get back to the relation between Problems 5.1 and 5.2.

Proposition 5.6. *Let C be a closed salient cone. Then for any choice of H and u there is some $\nu > 0$ such that (7) holds.*

Proof. After choosing H and u, choose $\nu > 0$ and a simplex cone S with
$$\nu{\uparrow}C \subseteq S \subseteq C.$$

We then have

$$(\nu\!\uparrow\! C)^{\max} \subseteq S^{\max} = S^{\min} \subseteq C^{\min}. \qquad\qquad \square$$

For any inclusion of cones $C \subseteq D$, we thus also have $(\nu\uparrow C)^{\max} \subseteq D^{\min}$. By suitable choice of u, we can also find a uniform bound on ν that only depends on the dimension:

Theorem 5.7. *Let $C \subseteq \mathbb{R}^d$ be a closed salient cone. Then for any choice of H, there is an order unit $u \in C$ such that the inclusion $(\nu\uparrow C)^{\max} \subseteq C^{\min}$ holds with $\nu = 1/(d+1)$.*

Proof. In the proof of Proposition 5.6, we apply the main theorem of [21]: whenever one inscribes into a convex body in $\mathbb{R}^{d-1}$ a simplex of maximal volume, then scaling the body with ratio $1/(d+1)$ from the barycenter of the simplex will make it contained in the simplex. $\qquad\square$

We can also recover the factor of inverse dimension from [7] in the presence of symmetry. Since we talk about cones as opposed to compact convex bodies, this dimension is our $d-1$:

Theorem 5.8. *Let $C \subseteq \mathbb{R}^d$ be a closed salient cone, and assume $C \cap (u + H)$ is symmetric with respect to u. Then the inclusion (7) holds with $\nu = 1/(d-1)$.*

Proof. As in the previous proof, one can use the simplex of maximal volume contained in the centrally symmetric convex body $C \cap (u + H)$. Grünbaum showed that the scaling factor can then be taken equal to the dimension [11, p. 259]. But since his method would not necessarily yield the center of symmetry u as the center of scaling, we argue slightly differently.

For notational simplicity, we assume $H = \mathbb{R}^{d-1} \times \{0\} \cong \mathbb{R}^{d-1}$ and $u = (0, \ldots, 0, 1)$, which we take to be the origin of $u + H$ as identified with $\mathbb{R}^{d-1}$. Set $\tilde{C} := C \cap (u + H) \subseteq \mathbb{R}^{d-1}$ and let $S \subseteq \tilde{C}$ be a simplex of maximal volume. Let b denote the barycenter of S. If F is a face of S and v its opposite vertex, then the translate of F through v, which is $F + \frac{d}{d-1}(v - b)$, is a hyperplane supporting $\tilde{C}$, since otherwise we could increase the volume of the simplex. The simplex defined by all these translates is $-(d-1)(S - b) + b$, and therefore $\tilde{C} \subseteq -(d-1)(S-b) + b$. Symmetry of $\tilde{C}$ then implies

$$\frac{1}{d-1}\tilde{C} + \frac{d}{d-1}b \subseteq S.$$

Now assume $(A_1, \ldots, A_d) \in C_s^{\max}$. Then

$$\left(\frac{1}{d-1}A_1 + \frac{d}{d-1}b_1 A_d, \ldots, \frac{1}{d-1}A_{d-1} + \frac{d}{d-1}b_{d-1}A_d, A_d \right) \in C_s^{\min},$$

by the argument used for Proposition 5.6. By symmetry we get the same result with $-b$ instead of b, and after adding and dividing by 2, we arrive at the desired conclusion,

$$\left(\frac{1}{d-1}A_1, \ldots, \frac{1}{d-1}A_{d-1}, A_d \right) \in C_s^{\min}. \qquad\qquad \square$$

Remark 5.9. Now given an instance of Problem 5.1, choose $\nu > 0$ such that (7) holds for $C = \mathcal{S}_1(M_1, \ldots, M_d)$. For convenience of notation, chose coordinates such that $u = (0, \ldots, 0, 1)$ and $H = \mathbb{R}^{d-1} \times \{0\}$ as in the previous proof. Then

$$\nu \uparrow C = \mathcal{S}_1(\nu^{-1} M_1, \ldots, \nu^{-1} M_{d-1}, M_d).$$

Now property (5.1) guarantees that if Problem 5.2 with

$$M_i' := \begin{cases} \nu^{-1} M_i & \text{for } i < d, \\ M_d & \text{for } i = d. \end{cases}$$

in place of the M_i has a negative solution, then so does the original Problem 5.1. So with this modification, Problem 5.2 is a relaxation rather than a strengthening of Problem 5.1. Intuitively speaking, the closer the scaling factor ν is to 1, the smaller the gap between Problem 5.1 and Problem 5.2 will be.

In other applications, such as the matrix cube problem [3], one is directly interested in the largest ν for which an inclusion of the form $\nu \uparrow \mathcal{S}_1(M_1, \ldots, M_d) \subseteq \mathcal{S}_1(N_1, \ldots, N_d)$ holds. In this situation, semidefinite strengthening as in Problem 5.2 provides a lower bound on the optimal ν. In the case of the matrix cube problem, Theorem 5.8 applies, and we may conclude that the strengthening of Problem 5.2 differs by a factor of at most $d-1$ from the actual optimal value. This bound is neither dominated by nor dominating over the error bound of Ben-Tal and Nemirovski [3].

Acknowledgments

References

[1] M. Argerami, S. Coskey, M. Kalantar, M. Kennedy, and M. Lupini, *The classification problem for finitely generated operator systems and spaces*, preprint. arXiv:1411.0512. ↑1

[2] W. B. Arveson, *Subalgebras of C^*-algebras*, Acta Math. **123** (1969), 141–224. ↑1

[3] A. Ben-Tal and A. Nemirovski, *On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty*, SIAM J. Optim. **12** (2002), no. 3, 811–833 (electronic). ↑1, 2, 14, 15, 17

[4] G. Blekherman, P. A. Parrilo, and R. R. Thomas (eds.), *Semidefinite optimization and convex algebraic geometry*, MOS-SIAM Series on Optimization, vol. 13, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2013. ↑1, 3

[5] M. D. Choi, *Completely positive linear maps on complex matrices*, Linear Algebra and Appl. **10** (1975), 285–290. ↑13, 14

[6] M. D. Choi and E. G. Effros, *Injectivity and operator spaces*, J. Functional Analysis **24** (1977), no. 2, 156–209. ↑2

[7] K. R. Davidson, A. Dor-On, O. M. Shalit, and B. Solel, *Dilations, inclusions of matrix convex sets, and completely positive maps*, preprint. arXiv:1601.07993. ↑1, 14, 15, 16

[8] K. Davidson and M. Kennedy, *The Choquet boundary of an operator system*, Duke Math. J. **164** (2015), no. 15, 2989–3004. arXiv:1303.3252. ↑1, 2

[9] E. G. Effros and S. Winkler, *Matrix convexity: operator analogues of the bipolar and Hahn-Banach theorems*, J. Funct. Anal. **144** (1997), no. 1, 117–152. ↑4

[10] E. Evert, J.W. Helton, I. Klep, and S. McCullough, *Extreme points of matrix convex sets, free spectrahedra and dilation theory*, in preparation. ↑13

[11] B. Grünbaum, *Measures of symmetry for convex sets*, Proc. Sympos. Pure Math., 1963, pp. 233–270. ↑16

[12] 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. arXiv:1412.1481. ↑1, 13, 14, 15

[13] J. W. Helton, I. Klep, and S. McCullough, *The matricial relaxation of a linear matrix inequality*, Math. Program. **138** (2013), no. 1-2, Ser. A, 401–445. arXiv:1003.0908. ↑1, 14

[14] J. W. Helton and S. McCullough, *Every convex free basic semi-algebraic set has an LMI representation*, Ann. of Math. (2) **176** (2012), no. 2, 979–1013. arXiv:0908.4352. ↑5

[15] F. John, *Extremum problems with inequalities as subsidiary conditions*, Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948, 1948, pp. 187–204. ↑

[16] K. Kellner, *Positivstellensatz Certificates for Containment of Polyhedra and Spectrahedra*, Doctoral Thesis, University of Frankfurt, Germany. http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/37816. ↑13

[17] K. Kellner and T. Theobald, *Sum of squares certificates for containment of $\mathcal{H}$-polytopes in $\mathcal{V}$-polytopes*, SIAM J. Discrete Math. **30** (2016), no. 2, 763–776. arXiv:1409.5008. ↑1

[18] K. Kellner, T. Theobald, and Ch. Trabandt, *A semidefinite hierarchy for containment of spectrahedra*, SIAM J. Optim. **25** (2015), no. 2, 1013–1033. arXiv:1308.5076. ↑1, 14, 15

[19] _____, *Containment problems for polytopes and spectrahedra*, SIAM J. Optim. **23** (2013), no. 2, 1000–1020. arXiv:1204.4313. ↑1, 2, 14

[20] V. Klee, *Some characterizations of convex polyhedra*, Acta Math. **102** (1959), 79–107. ↑7

[21] M. Lassak, *Approximation of convex bodies by inscribed simplices of maximum volume*, Beitr. Algebra Geom. **52** (2011), no. 2, 389–394. ↑16

[22] V. Paulsen, *Completely bounded maps and operator algebras*, Cambridge Studies in Advanced Mathematics, vol. 78, Cambridge University Press, Cambridge, 2002. ↑2

[23] V. Paulsen, I. Todorov, and M. Tomforde, *Operator system structures on ordered spaces*, Proc. Lond. Math. Soc. (3) **102** (2011), no. 1, 25–49. ↑6

[24] M. Ramana and A. J. Goldman, *Some geometric results in semidefinite programming*, J. Global Optim. **7** (1995), no. 1, 33–50. ↑10

[25] J. F. Sturm, *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*, Optimization Methods and Software **11–12** (1999), 625–653. Software available at https://github.com/sqlp/sedumi. ↑3

[26] H. Wolkowicz and R. Saigal and L. Vandenberghe (ed.), *Handbook of semidefinite programming*, International Series in Operations Research & Management Science, 27, Kluwer Academic Publishers, Boston, MA, 2000. Theory, algorithms, and applications. ↑3

[27] G. M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics, vol. 152, Springer-Verlag, New York, 1995. ↑12

TOBIAS FRITZ, MAX PLANCK INSTITUTE FOR MATHEMATICS IN THE SCIENCES, LEIPZIG, GERMANY
E-mail address: `fritz@mis.mpg.de`

TIM NETZER, UNIVERSITÄT INNSBRUCK, AUSTRIA
E-mail address: `tim.netzer@uibk.ac.at`

ANDREAS THOM, TU DRESDEN, GERMANY
E-mail address: `andreas.thom@tu-dresden.de`