A spectrahedral representation of the first derivative relaxation of the positive semidefinite cone

James Saunderson*

September 5, 2018

Abstract

If X is an $n \times n$ symmetric matrix, then the directional derivative of $X \mapsto \det(X)$ in the direction I is the elementary symmetric polynomial of degree n-1 in the eigenvalues of X. This is a polynomial in the entries of X with the property that it is hyperbolic with respect to the direction I. The corresponding hyperbolicity cone is a relaxation of the positive semidefinite (PSD) cone known as the first derivative relaxation (or Renegar derivative) of the PSD cone. A spectrahedal cone is a convex cone that has a representation as the intersection of a subspace with the cone of PSD matrices in some dimension. We show that the first derivative relaxation of the PSD cone is a spectrahedral cone, and give an explicit spectrahedral description of size $\binom{n+1}{2}-1$. The construction provides a new explicit example of a hyperbolicity cone that is also a spectrahedron. This is consistent with the generalized Lax conjecture, which conjectures that every hyperbolicity cone is a spectrahedron.

1 Introduction

1.1 Preliminaries

Hyperbolic polynomials, hyperbolicity cones, and spectrahedra A multivariate polynomial p, homogeneous of degree d in n variables, is hyperbolic with respect to $e \in \mathbb{R}^n$ if $p(e) \neq 0$ and for all x, the univariate polynomial $t \mapsto p(x - te)$ has only real roots. Associated with such a polynomial is a cone

$$\Lambda_+(p,e) = \{x \in \mathbb{R}^n : \text{all roots of } t \mapsto p(x-te) \text{ are non-negative} \}.$$

A foundational result of Gårding [Går59] is that $\Lambda_+(p,e)$ is actually a convex cone, called the *closed* hyperbolicity cone associated with p and e.

For example $p(x) = \prod_{i=1}^n x_i$ is hyperbolic with respect to 1_n , the vector of all ones, and the corresponding closed hyperbolicity cone is the non-negative orthant, \mathbb{R}^n_+ . Similarly $p(X) = \det(X)$ (where X is a symmetric $n \times n$ matrix), is hyperbolic with respect to the identity matrix I, and the corresponding closed hyperbolicity cone is the positive semidefinite cone \mathcal{S}^n_+ .

If a polynomial p has a representation of the form

$$p(x) = \det\left(\sum_{i=1}^{n} A_i x_i\right) \tag{1}$$

^{*}Department of Electrical and Computer Systems Engineering, Monash University, VIC 3800, Australia. Email: james.saunderson@monash.edu

for symmetric matrices A_1, \ldots, A_n , and there exists $e \in \mathbb{R}^n$ such that $\sum_{i=1}^n A_i e_i$ is positive definite, we say that p has a definite determinantal representation. In this case p is hyperbolic with respect to e. The associated closed hyperbolicity cone is

$$K = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n A_i x_i \succeq 0 \right\}$$
 (2)

where we write $X \succeq 0$ to indicate that X is positive semidefinite (and $X \succ 0$ to indicate that X is positive definite). Such convex cones are called *spectrahedral cones*. If the matrices A_1, A_2, \ldots, A_n are $d \times d$ we call (2) a *spectrahedral representation of size* d.

Derivative relaxations One way to produce new hyperbolic polynomials is to take directional derivatives of hyperbolic polynomials in directions of hyperbolicity [ABG70, Section 3.10], a construction emphasized in the context of optimization by Renegar [Ren06]. If p has degree d and is hyperbolic with respect to e, then for $k = 0, 1, \ldots, d$, the kth directional derivative in the direction e, i.e.,

$$D_e^{(k)}p(x) = \left. \frac{d^k}{dt^k} p(x+te) \right|_{t=0},$$

is also hyperbolic with respect to e. Moreover

$$\Lambda_+(D_e^{(k)}p,e) \supseteq \Lambda_+(D_e^{(k-1)}p,e) \supseteq \cdots \supseteq \Lambda_+(p,e)$$

so the hyperbolicity cones of the directional derivatives form a sequence of *relaxations* of the original hyperbolicity cone.

• Suppose $p(x) = \prod_{i=1}^{n} x_i$ and $e = 1_n$. Then, for $k = 0, 1, \dots, n$,

$$D_{1n}^{(k)}p(x) = k!e_{n-k}(x)$$

where e_{n-k} is the elementary symmetric polynomial of degree n-k in n variables. We use the notation $\mathbb{R}^{n,(k)}_+$ for $\Lambda_+(e_{n-k},1_n)$, the closed hyperbolicity cone corresponding to e_{n-k} .

• Suppose $p(X) = \det(X)$ is the determinant restricted to $n \times n$ symmetric matrices, and $e = I_n$ is the $n \times n$ identity matrix. Then, for $k = 0, 1, \ldots, n$,

$$D_{l_n}^{(k)}p(X) = k! E_{n-k}(X) = k! e_{n-k}(\lambda(X))$$

where $E_{n-k}(X)$ is the elementary symmetric polynomial of degree n-k in the eigenvalues of X or, equivalently, the coefficient of t^k in $\det(X+tI_n)$. We use the notation $\mathcal{S}^{n,(k)}_+$ for $\Lambda_+(E_{n-k},I_n)$, the closed hyperbolicity cone corresponding to E_{n-k} . We use the notation $\lambda(X)$ for the eigenvalues of a symmetric matrix X ordered so that $|\lambda_1(X)| \geq |\lambda_2(X)| \geq \cdots \geq |\lambda_n(X)|$. We use this order so that $\lambda_i(X^2) = \lambda_i(X)^2$ for all i.

The focus of this paper is the cone $\mathcal{S}_{+}^{n,(1)}$, the hyperbolicity cone associated with E_{n-1} . In particular, we consider whether $\mathcal{S}_{+}^{n,(1)}$ can be expressed as a 'slice' of some higher dimensional positive semidefinite cone. Such a description allows one to reformulate hyperbolic programs with respect to $\mathcal{S}_{+}^{n,(1)}$ (linear optimization over affine 'slices' of $\mathcal{S}_{+}^{n,(1)}$) as semidefinite programs.

Generalized Lax conjecture We have seen that every spectrahedral cone is a closed hyperbolicity cone. The generalized Lax conjecture asks whether the converse holds, i.e., whether every closed hyperbolicity cone is also a spectrahedral cone. The original Lax conjecture, now a theorem due to Helton and Vinnikov [HV07] (see also [LPR05]), states that if p is a trivariate polynomial, homogeneous of degree d, and hyperbolic with respect to $e \in \mathbb{R}^3$, then p has a definite determinantal representation. While a direct generalization of this algebraic result does not hold in higher dimensions [Brä11], the following geometric conjecture remains open.

Conjecture 1 (Generalized Lax Conjecture (geometric version)). Every closed hyperbolicity cone is spectrahedral.

An equivalent algebraic formulation of this conjecture is as follows.

Conjecture 2 (Generalized Lax Conjecture (algebraic version)). If p is hyperbolic with respect to $e \in \mathbb{R}^n$, then there exists a polynomial q, hyperbolic with respect to $e \in \mathbb{R}^n$, such that qp has a definite determinantal representation and $\Lambda_+(q,e) \supseteq \Lambda_+(p,e)$.

The algebraic version of the conjecture implies the geometric version because it implies the existence of a multiplier q such that the hyperbolicity cone associated with qp is spectrahedral and $\Lambda_+(qp,e) = \Lambda_+(p,e) \cap \Lambda_+(q,e) = \Lambda_+(p,e)$. To see that the geometric version implies the algebraic version requires more algebraic machinery, and is discussed, for instance, in [Vin12, Section 2].

1.2 Main result: a spectrahedral representation of $\mathcal{S}^{n,(1)}_+$

In this paper, we show that $\mathcal{S}^{n,(1)}_+$, the first derivative relaxation of the positive semidefinite cone, is spectrahedral. We give an explicit spectrahedral representation of $\mathcal{S}^{n,(1)}_+$ (see Theorem 1 to follow). Moreover, in Theorem 3 in Section 2 we find an explicit hyperbolic polynomial q such that $q(X)E_{n-1}(X)$ has a definite determinantal representation and $\Lambda_+(q,I) \supseteq \mathcal{S}^{n,(1)}_+$.

Theorem 1. Let $d = \binom{n+1}{2} - 1$ and let B_1, \ldots, B_d be any basis for the d-dimensional space of real symmetric $n \times n$ matrices with trace zero. If $\mathcal{B}(X)$ is the $d \times d$ symmetric matrix with i, j entry equal to $\operatorname{tr}(B_i X B_j)$ then

$$\mathcal{S}_{+}^{n,(1)} = \{ X \in \mathcal{S}^n : \mathcal{B}(X) \succeq 0 \}. \tag{3}$$

Section 2 is devoted to the proof of this result. At this stage we make a few remarks about the statement and some of its consequences.

- The spectrahedral representation of $\mathcal{S}_{+}^{n,(1)}$ in Theorem 1 has size $d = \binom{n+1}{2} 1 = \frac{1}{2}(n+2)(n-1)$. This is about half the size of the smallest previously known *projected* spectrahedral representation of $\mathcal{S}_{+}^{n,(1)}$, i.e., representation as the image of a spectrahedral cone under a linear map [SP15].
- A straightforward extension of this result shows that if p has a definite determinantal representation and e is a direction of hyperbolicity for p, then the hyperbolicity cone associated with the directional derivative $D_e p$ is spectrahedral. We discuss this in Section 3.1.
- It also follows from Theorem 1 that $\mathbb{R}^{n,(2)}_+$, the second derivative relaxation of the orthant in the direction 1_n , has a spectrahedral representation of size $\binom{n}{2} 1$. We discuss this in Section 3.1. This representation is significantly smaller than the size $O(n^{n-3})$ representation constructed by Brändén [Brä14], and about half the size of the smallest previously known projected spectrahedral representation of $\mathbb{R}^{n,(2)}_+$ [SP15].

1.3 Related work

We briefly summarize related work on spectrahedral and projected spectrahedral representations of the hyperbolicity cones $\mathbb{R}^{n,(k)}_+$ and $\mathcal{S}^{n,(k)}_+$. Sanyal [San13] showed that $\mathbb{R}^{n,(1)}_+$ is spectrahedral by giving the following explicit definite determinantal representation of $e_{n-1}(x)$, which we use repeatedly in the paper.

Proposition 1. If $1_n^{\perp} = \{x \in \mathbb{R}^n : 1_n^T x = 0\}$, and V_n is a $n \times (n-1)$ matrix with columns spanning 1_n^{\perp} , then there is a positive constant c such that

$$c e_{n-1}(x) = \det(V_n^T \operatorname{diag}(x) V_n)$$
 and so $\mathbb{R}^{n,(1)}_+ = \{x \in \mathbb{R}^n : V_n^T \operatorname{diag}(x) V_n \succeq 0\}.$

This representation is also implicit in the work of Choe, Oxley, Sokal, and Wagner [COSW04]. Zinchenko [Zin08], gave a projected spectrahedral representation of $\mathbb{R}^{n,(1)}_+$. Brändén [Brä14], established that each of the cones $\mathbb{R}^{n,(k)}_+$ are spectrahedral by constructing graphs G with edges weighted by linear forms in x, such that the edge weighted Laplacian $L_G(x)$ is positive semidefinite if and only if $x \in \mathbb{R}^{n,(k)}_+$. Amini showed that the hyperbolicity cones associated with certain multivariate matching polynomials are spectrahedral [Ami16], and used these to find new spectrahedral representations of the cones $\mathbb{R}^{n,(k)}_+$ of size $\frac{(n-1)!}{(k-1)!} + 1$.

Explicit projected spectrahedral representations of the cones $\mathcal{S}_{+}^{n,(k)}$ of size $O(n^2 \min\{k, n-k\})$ were given by Saunderson and Parrilo [SP15], leaving open (except in the cases k=n-2, n-1) the question of whether these cones are spectrahedra. The main result of this paper is that $\mathcal{S}_{+}^{n,(1)}$ is a spectrahedron.

2 Proof of Theorem 1

In this section we give two proofs of Theorem 1. The first proof is convex geometric in nature whereas the second is algebraic in nature. Both arguments are self-contained. We present the geometric argument first because it suggests the choice of multiplier q for the algebraic argument.

Both arguments take advantage of the fact that the cone $\mathcal{S}_{+}^{n,(1)}$ satisfies $Q\mathcal{S}_{+}^{n,(1)}Q^{T}=\mathcal{S}_{+}^{n,(1)}$ for all $Q\in O(n)$. One way to see this is to observe that the hyperbolic polynomial $E_{n-1}(X)$ that determines the cone satisfies $E_{n-1}(QXQ^{T})=E_{n-1}(X)$ for all $Q\in O(n)$ and the direction of hyperbolicity (the identity) is also invariant under this group action.

2.1 Geometric argument

We begin by stating a slight reformulation of Sanyal's spectrahedral representation (Proposition 1).

Proposition 2. Let $1_n^{\perp} = \{y \in \mathbb{R}^n : 1_n^T y = 0\}$ be the subspace of \mathbb{R}^n orthogonal to 1_n . Then

$$\mathbb{R}^{n,(1)}_+ = \{x \in \mathbb{R}^n \ : \ y^T \operatorname{diag}(x) y \geq 0 \ \text{ for all } y \in 1^\perp_n \}.$$

Proof. This follows from Proposition 1 since $V_n^T \operatorname{diag}(x) V_n \succeq 0$ holds if and only if $u^T V_n^T \operatorname{diag}(x) V_n u \geq 0$ for all $u \in \mathbb{R}^{n-1}$ which holds if and only if $y^T \operatorname{diag}(x) y \geq 0$ for all $y \in \mathbb{1}_n^{\perp}$.

In this section we establish a 'matrix' analogue of Proposition 2.

Theorem 2. Let $I_n^{\perp} = \{Y \in \mathcal{S}^n : \operatorname{tr}(Y) = 0\}$ be the subspace of $n \times n$ symmetric matrices with trace zero. Then

$$\mathcal{S}_{+}^{n,(1)} = \{ X \in \mathcal{S}^n : \operatorname{tr}(YXY) \ge 0, \text{ for all } Y \in I_n^{\perp} \}.$$
 (4)

The concrete spectrahedral description given in Theorem 1 follows immediately from Theorem 2. Indeed if B_1, B_2, \ldots, B_d are a basis for I_n^{\perp} then an arbitrary $Y \in I_n^{\perp}$ can be written as $Y = \sum_{i=1}^d y_i B_i$. The condition $\operatorname{tr}(YXY) \geq 0$ for all $Y \in I_n^{\perp}$ is equivalent to

$$\sum_{i,j=1}^d y_i y_j \operatorname{tr}(B_i X B_j) \ge 0 \text{ for all } y \in \mathbb{R}^d \text{ which holds if and only if } \mathcal{B}(X) \succeq 0.$$

of Theorem 2. The convex cone $\mathcal{S}^{n,(1)}_+$ is invariant under the action of the orthogonal group on $n \times n$ symmetric matrices by congruence transformations. Similarly, the convex cone

$$\{X \in \mathcal{S}^n : \operatorname{tr}(YXY) \ge 0 \text{ for all } Y \in I_n^{\perp}\}$$

is invariant under the same action of the orthogonal group. This is because $X \in I_n^{\perp}$ if and only if $QXQ^T \in I^{\perp}$ for any orthogonal matrix Q.

Because of these invariance properties, the following (straightforward) result tells us that we can establish Theorem 2 by showing that the diagonal 'slices' of these two convex cones agree.

Lemma 1. Let $K_1, K_2 \subset \mathcal{S}^n$ be such that $QK_1Q^T = K_1$ for all $Q \in O(n)$ and $QK_2Q^T = K_2$ for all $Q \in O(n)$. If $\{x \in \mathbb{R}^n : \operatorname{diag}(x) \in K_1\} = \{x \in \mathbb{R}^n : \operatorname{diag}(x) \in K_2\}$ then $K_1 = K_2$.

Proof. Assume that $X \in K_1$. Then there exists Q such that $QXQ^T = \operatorname{diag}(\lambda(X))$. Since K_1 is invariant under orthogonal congruence, $\operatorname{diag}(\lambda(X)) \in K_1$. By assumption, it follows that $\operatorname{diag}(\lambda(X)) \in K_2$. Since K_2 is invariant under orthogonal congruence, $X = Q^T \operatorname{diag}(\lambda(X))Q \in K_2$. This establishes that $K_1 \subseteq K_2$. Reversing the roles of K_1 and K_2 completes the argument.

Relating the diagonal slices To complete the proof of Theorem 2, it suffices (by Lemma 1) to show that the diagonal slices of the left- and right-hand sides of (4) are equal. Since the diagonal slice of $\mathcal{S}_{+}^{n,(1)}$ is $\mathbb{R}_{+}^{n,(1)}$, it is enough (by Proposition 2) to establish the following result.

Lemma 2.

$$\{x \in \mathbb{R}^n \ : \ \operatorname{tr}(Y\operatorname{diag}(x)Y) \geq 0 \ \text{ for all } Y \in I_n^{\perp} \} = \\ \{x \in \mathbb{R}^n \ : \ y^T\operatorname{diag}(x)y \geq 0 \ \text{ for all } y \in \mathbb{1}_n^{\perp} \}.$$

Proof. Suppose that $\operatorname{tr}(Y\operatorname{diag}(x)Y) \geq 0$ for all $Y \in I_n^{\perp}$. Let $y \in I_n^{\perp}$. Then $\operatorname{diag}(y) \in I_n^{\perp}$ and so it follows that $\operatorname{tr}(\operatorname{diag}(y)\operatorname{diag}(x)\operatorname{diag}(y)) = y^T\operatorname{diag}(x)y \geq 0$. This shows that the left hand side is a subset of the right hand side.

For the reverse inclusion suppose that $y^T \operatorname{diag}(x)y \geq 0$ for all $y \in 1_n^{\perp}$. Let $Y \in I_n^{\perp}$. Suppose the symmetric group on n symbols, S_n , acts on \mathbb{R}^n by permutations. Then for every $\sigma \in S_n$, we have that $\sigma \cdot \lambda(Y) \in 1_n^{\perp}$ and thus

$$\operatorname{tr}(\operatorname{diag}(\sigma \cdot \lambda(Y^2))\operatorname{diag}(x)) = (\sigma \cdot \lambda(Y))^T\operatorname{diag}(x)(\sigma \cdot \lambda(Y)) \ge 0.$$

(Here we have used $\lambda_i(Y^2) = \lambda_i(Y)^2$, by our definition of $\lambda(\cdot)$.)

The diagonal of a symmetric matrix is a convex combination of permutations of its eigenvalues, a result due to Schur [Sch23] (see also, e.g., [MOA79]). Hence diag(Y^2) is a convex combination of permutations of $\lambda(Y^2)$, i.e.,

$$\operatorname{diag}(Y^2) = \sum_{\sigma \in S_n} \eta_{\sigma} \left(\sigma \cdot \lambda(Y^2) \right)$$

where the η_{σ} satisfy $\eta_{\sigma} \geq 0$ and $\sum_{\sigma \in S_n} \eta_{\sigma} = 1$. It then follows that

$$\operatorname{tr}(Y\operatorname{diag}(x)Y) = \operatorname{tr}(\operatorname{diag}(Y^2)\operatorname{diag}(x)) = \sum_{\sigma \in S_n} \eta_\sigma \operatorname{tr}(\operatorname{diag}(\sigma \cdot \lambda(Y^2))\operatorname{diag}(x)) \geq 0.$$

This shows that the right hand side is a subset of the left hand side.

This completes the proof of Theorem 2.

2.2 Algebraic argument

In this section, we establish the following algebraic version of Theorem 1.

Theorem 3. Let $n \geq 2$ and B_1, \ldots, B_d be a basis for I_n^{\perp} , the subspace of $n \times n$ symmetric matrices with trace zero. Then there is a positive constant c (depending on the choice of basis) such that

- 1. $q(X) = \prod_{1 \le i \le j \le n} (\lambda_i(X) + \lambda_j(X))$ is hyperbolic with respect to I_n ;
- 2. the hyperbolicity cone associated with q satisfies

$$\Lambda_{+}(q, I_n) = \{ X \in \mathcal{S}^n : \lambda_i(X) + \lambda_j(X) \ge 0 \text{ for all } 1 \le i < j \le n \} \supseteq \mathcal{S}_{+}^{n,(1)};$$

3. $q(X)E_{n-1}(X)$ has a definite determinantal representation as

$$c q(X)E_{n-1}(X) = \det(\mathcal{B}(X)).$$

We remark that q(X) is defined as a symmetric polynomial in the eigenvalues of X, and so can be expressed as a polynomial in the entries of X. Although our argument does not use this fact, it can be shown that $q(X) = \det(\mathcal{L}_2(X))$ where $\mathcal{L}_2(X)$ is the second additive compound matrix of X [Fie74]. This means that q is not only hyperbolic with respect to I_n , but also has a definite determinantal representation.

of Theorem 3. The three items in the statement of Theorem 3 are established in the following three Lemmas (Lemmas 3, 4, and 5).

Lemma 3. If $q(X) = \prod_{1 \le i \le j \le n} (\lambda_i(X) + \lambda_j(X))$ then q is hyperbolic with respect to I_n .

Proof. First observe that $q(I_n) = 2^{\binom{n}{2}} \neq 0$. Moreover, for any real t,

$$q(X - tI_n) = \prod_{1 \le i < j \le n} (\lambda_i(X - tI_n) + \lambda_j(X - tI_n)) = \prod_{1 \le i < j \le n} (\lambda_i(X) + \lambda_j(X) - 2t)$$

which has $\binom{n}{2}$ real roots given by $\frac{1}{2}(\lambda_i(X) + \lambda_j(X))$ for $1 \le i < j \le n$. Hence q is hyperbolic with respect to I_n .

Lemma 4. If $n \geq 2$ then

$$\Lambda_{+}(q, I_n) = \{ X \in \mathcal{S}^n : \lambda_i(X) + \lambda_j(X) \ge 0 \text{ for all } 1 \le i < j \le n \} \supseteq \mathcal{S}_{+}^{n,(1)}.$$

Proof. Since the roots of $t \mapsto q(X - tI_n)$ are $\frac{1}{2}(\lambda_i(X) + \lambda_j(X))$, the description of $\Lambda_+(q, I_n)$ is immediate. Both sides of the inclusion are invariant under congruence by orthogonal matrices. By Lemma 1 it is enough to show that the inclusion holds for the diagonal slices of both sides. Note that

$$\{x \in \mathbb{R}^n : \operatorname{diag}(x) \in \Lambda_+(q, I_n)\} = \{x \in \mathbb{R}^n : x_i + x_j \ge 0 \text{ for all } 1 \le i < j \le n\}.$$

Hence it is enough to establish that

$$\{x \in \mathbb{R}^n : x_i + x_j \ge 0 \text{ for all } 1 \le i < j \le n\} \supseteq \mathbb{R}^{n,(1)}_+.$$
 (5)

To do so, we use the characterization of $\mathbb{R}^{n,(1)}_+$ from Proposition 2. This tells us that if $x \in \mathbb{R}^{n,(1)}_+$ then $v^T \operatorname{diag}(x)v = \sum_{\ell=1}^n x_\ell v_\ell^2 \geq 0$ for all $v \in 1_n^\perp$. In particular, let v be the element of 1_n^\perp with $v_i = 1$ and $v_j = -1$ and $v_k = 0$ for $k \notin \{i, j\}$. Then, if $x \in \mathbb{R}^{n,(1)}_+$ it follows that $\sum_{\ell=1}^n x_\ell v_\ell^2 = x_i + x_j \geq 0$. This completes the proof.

Lemma 5. If B_1, \ldots, B_d is a basis for I_n^{\perp} , then there is a positive constant c (depending on the choice of basis) such that

$$c q(X)E_{n-1}(X) = \det(\mathcal{B}(X)).$$

Proof. Since both sides are invariant under orthogonal congruence, it is enough to show that the identity holds for diagonal matrices. In other words, it is enough to show that

$$c \prod_{1 \le i < j \le n} (x_i + x_j) e_{n-1}(x) = \det(\mathcal{B}(\operatorname{diag}(x))).$$

Since a change of basis for the subspace of symmetric matrices with trace zero only changes $\det(\mathcal{B}(X))$ by a positive constant (which is one if the change of basis is orthogonal with respect to the trace inner product), it is enough to choose a particular basis for the subspace of symmetric matrices with trace zero, and show that the identity holds for a particular constant.

Let $v_1, v_2, \ldots, v_{n-1}$ be a basis for $1_n^{\perp} = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 0\}$. Let M_{ij} be the $n \times n$ matrix with a one in the (i,j) and the (j,i) entry, and zeros elsewhere. Clearly the M_{ij} for $1 \leq i < j \leq n$ form a basis for the subspace of symmetric matrices with zero diagonal. Together $\operatorname{diag}(v_1), \operatorname{diag}(v_2), \ldots, \operatorname{diag}(v_{n-1})$ and M_{ij} for $1 \leq i < j \leq n$ form a basis for the subspace of symmetric matrices with trace zero.

Using this basis we evaluate the matrix $\mathcal{B}(\operatorname{diag}(x))$. We note that

$$\operatorname{tr}(\operatorname{diag}(v_i)\operatorname{diag}(x)\operatorname{diag}(v_j)) = v_i^T\operatorname{diag}(x)v_j \quad \text{for } 1 \leq i, j \leq n$$

 $\operatorname{tr}(\operatorname{diag}(v_i)\operatorname{diag}(x)M_{jk}) = 0 \quad \text{for all } 1 \leq i \leq n \text{ and } 1 \leq j < k \leq n$

since M_{ik} has zero diagonal, and that

$$\operatorname{tr}(M_{ij}\operatorname{diag}(x)M_{k\ell}) = \begin{cases} x_i + x_j & \text{if } i = k \text{ and } j = \ell \\ 0 & \text{otherwise} \end{cases}$$

for all $1 \le i < j \le n$ and $1 \le k < \ell \le n$. This means that $\mathcal{B}(\operatorname{diag}(x))$ is block diagonal, and so

$$\det(\mathcal{B}(\operatorname{diag}(x))) = \prod_{1 \le i < j \le n} (x_i + x_j) \det(V_n^T \operatorname{diag}(x) V_n)$$
(6)

where V_n is the $n \times (n-1)$ matrix with columns v_1, v_2, \ldots, v_n . By Proposition 1, there is a positive constant c such that

$$\det(V_n^T \operatorname{diag}(x)V_n) = c e_{n-1}(x), \tag{7}$$

Combining (6) and (7) gives the stated result.

This completes the proof of Theorem 3. \Box

3 Discussion

3.1 Consequences of Theorem 1

A straightforward consequence of Theorem 1 is that if p has a definite determinantal representation, and e is a direction of hyperbolicity for p, then the hyperbolicity cone associated with the directional derivative $D_e p$ is spectrahedral.

Corollary 1. If $p(x) = \det(\sum_{i=1}^{n} A_i x_i)$ for symmetric $\ell \times \ell$ matrices A_1, \ldots, A_n , and $A_0 = \sum_{i=1}^{m} A_i e_i$ is positive definite, then $\Lambda_+(D_e p, e)$ has a spectrahedral representation of size $\binom{\ell+1}{2} - 1$.

Proof. The hyperbolicity cone $\Lambda_{+}(D_{e}p, e)$ can be expressed as

$$\Lambda_{+}(D_{e}p, e) = \left\{ x \in \mathbb{R}^{n} : \sum_{i=1}^{n} A_{0}^{-1/2} A_{i} A_{0}^{-1/2} x_{i} \in \mathcal{S}_{+}^{n,(1)} \right\}.$$

(see, e.g., [SP15, Proposition 4]). Applying Theorem 1 then gives

$$\Lambda_{+}(D_{e}p, e) = \left\{ x \in \mathbb{R}^{n} : \mathcal{B}\left(\sum_{i=1}^{n} A_{0}^{-1/2} A_{i} A_{0}^{-1/2} x_{i}\right) \succeq 0 \right\}.$$

Our main result also yields a spectrahedral representation of $\mathbb{R}^{n,(2)}_+$, the second derivative relaxation of the non-negative orthant, of size $\binom{n}{2}-1$. This is, in fact, a special case of Corollary 1. In the statement below, V_n is any $n \times (n-1)$ matrix with columns that span 1^{\perp}_n .

Corollary 2. The hyperbolicity cone $\mathbb{R}^{n,(2)}_+$ has a spectrahedral representation of size $\binom{n}{2} - 1$ given by

$$\mathbb{R}_{+}^{n,(2)} = \{ x \in \mathbb{R}^n : \mathcal{B}(V_n^T \operatorname{diag}(x)V_n) \succeq 0 \}.$$

Proof. First, we use the fact that $\mathbb{R}^{n,(2)}_+ = \Lambda_+(D_{1n}e_{n-1},1_n)$. Then, by Sanyal's result (Proposition 1), we know that $e_{n-1}(x)$ has a definite determinantal representation. The stated result then follows directly from Corollary 1 with polynomial $p=e_{n-1}$ and direction $e=1_n$.

3.2 Questions

Constructing spectrahedral representations It is natural to ask for which values of k the cones $\mathcal{S}_{+}^{n,(k)}$ are spectrahedral. Our main result shows that $\mathcal{S}_{+}^{n,(1)}$ has a spectrahedral representation of size $d = \binom{n+1}{2} - 1$. The only other cases for which spectrahedral representations are known are the straightforward cases k = n - 1 and k = n - 2. If k = n - 1 then

$$\mathcal{S}_{+}^{n,(n-1)} = \{ X \in \mathcal{S}^n : \operatorname{tr}(X) \ge 0 \}$$

is a spectrahedron (with a representation of size 1). Since $\mathcal{S}_{+}^{n,(n-2)}$ is a quadratic cone, it is a spectrahedron. To give an explicit representation, let $d = \binom{n+1}{2} - 1$ and B_1, B_2, \ldots, B_d be an *orthonormal* basis (with respect to the trace inner product) for the subspace I_n^{\perp} . Now $X \in \mathcal{S}_{+}^{n,(n-2)}$ if and only if (see, e.g., [SP15, Section 5.1])

$$\operatorname{tr}(X) \ge 0 \text{ and } \operatorname{tr}(X)^2 - \operatorname{tr}(X^2) = \left[\sqrt{\frac{n-1}{n}}\operatorname{tr}(X)\right]^2 - \sum_{i=1}^d \operatorname{tr}(B_i X)^2 \ge 0.$$
 (8)

By a well-known spectrahedral representation of the second-order cone, (8) holds if and only if

$$\sqrt{\frac{n-1}{n}} \operatorname{tr}(X) I_d + \begin{bmatrix} \operatorname{tr}(B_1 X) & \operatorname{tr}(B_2 X) & \operatorname{tr}(B_3 X) & \cdots & \operatorname{tr}(B_d X) \\ \operatorname{tr}(B_2 X) & -\operatorname{tr}(B_1 X) & 0 & \cdots & 0 \\ \operatorname{tr}(B_3 X) & 0 & -\operatorname{tr}(B_1 X) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \operatorname{tr}(B_d X) & 0 & 0 & \cdots & -\operatorname{tr}(B_1 X) \end{bmatrix} \succeq 0.$$
(9)

So we see that $\mathcal{S}_{+}^{n,(n-2)}$ has a spectrahedral representation of size $d = \binom{n+1}{2} - 1$. At this stage, it is unclear how to extend the approach in this paper to the remaining cases.

Question 1. Are the cones $S_{+}^{n,(k)}$ spectrahedral for k = 2, 3, ..., n - 3?

At first glance, it may seem that Corollary 1 allows us to construct a spectrahedral representation for $\mathcal{S}^{n,(2)}_+$ from a spectrahedral representation for $\mathcal{S}^{n,(1)}_+$. However, this is not the case. To apply Corollary 1 to this situation, we would need a definite determinantal representation of $E_{n-1}(X)$, which our main result (Theorem 1) does not provide.

Lower bounds on size Another natural question concerns the size of spectrahedral representations of hyperbolicity cones. Given a hyperbolicity cone K, there is a unique (up to scaling) hyperbolic polynomial p of smallest degree d that vanishes on the boundary of K (see, e.g., [Kum16]). Clearly any spectrahedral representation must have size at least d, but it seems that in some cases the smallest spectrahedral representation (if it exists at all) must have larger size.

Question 2. Is there a spectrahedral representation of
$$S_+^{n,(1)}$$
 with size smaller than $\binom{n+1}{2} - 1$?

Recently, there has been considerable interest in developing methods for producing lower bounds on the size of projected spectrahedral descriptions of convex sets (see, e.g., [FGP⁺15]) . There has been much less development in the case of lower bounds on the size of spectrahedral descriptions. The main work in this direction is due to Kummer [Kum16]. For instance it follows from [Kum16, Theorem 1] that any spectrahedral representation of the quadratic cone $\mathcal{S}_{+}^{n,(n-2)}$ must have size at least $\frac{1}{2} \left[\binom{n+1}{2} - 1 \right]$. Furthermore, in the special case that $\binom{n+1}{2} - 1 = 2^k + 1$ for some k (which occurs if n=3 and k=2 or n=4 and k=3) then Kummer's work shows that any spectrahedral representation of $\mathcal{S}_{+}^{n,(n-2)}$ must have size at least $\binom{n+1}{2} - 1$. This establishes that the construction in (9) is optimal when n=3 and n=4. Furthermore, in the case n=3 we have that $\mathcal{S}_{+}^{n,(1)} = \mathcal{S}_{+}^{n,(n-2)}$. Hence our spectrahedral representation for $\mathcal{S}_{+}^{n,(1)}$ is also optimal if n=3.

Acknowledgments

I would like to thank Hamza Fawzi for providing very helpful feedback on a draft of this paper.

References

- [ABG70] M. F. Atiyah, R. Bott, and L. Gårding. Lacunas for hyperbolic differential operators with constant coefficients I. *Acta Mathematica*, 124(1):109–189, 1970. 2
- [Ami16] N. Amini. Spectrahedrality of hyperbolicity cones of multivariate matching polynomials. arXiv preprint arXiv:1611.06104, 2016. 4

- [Brä11] P. Brändén. Obstructions to determinantal representability. Advances in Mathematics, 226(2):1202–1212, 2011. 3
- [Brä14] P. Brändén. Hyperbolicity cones of elementary symmetric polynomials are spectrahedral. *Optimization Letters*, 8(5):1773–1782, 2014. 3, 4
- [COSW04] Y.-B. Choe, J. G. Oxley, A. D. Sokal, and D. G. Wagner. Homogeneous multivariate polynomials with the half-plane property. *Advances in Applied Mathematics*, 32(1):88–187, 2004. 4
- [FGP⁺15] H. Fawzi, J. Gouveia, P. A. Parrilo, R. Z. Robinson, and R. R. Thomas. Positive semidefinite rank. *Mathematical Programming*, 153(1):133–177, 2015. 9
- [Fie74] M. Fiedler. Additive compound matrices and an inequality for eigenvalues of symmetric stochastic matrices. *Czechoslovak Mathematical Journal*, 24(3):392–402, 1974. 6
- [Går59] L. Gårding. An inequality for hyperbolic polynomials. *Journal of Mathematics and Mechanics*, 8(6):957–965, 1959. 1
- [HV07] J. W. Helton and V. Vinnikov. Linear matrix inequality representation of sets. Communications in Pure and Applied Mathematics, 60(5):654–674, 2007. 3
- [Kum16] M. Kummer. Two results on the size of spectrahedral descriptions. SIAM Journal on Optimization, 26(1):589–601, 2016. 9
- [LPR05] A. S. Lewis, P. A. Parrilo, and M. V. Ramana. The Lax conjecture is true. *Proceedings* of the American Mathematical Society, 133(9):2495–2500, 2005. 3
- [MOA79] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: theory of majorization and its applications*, volume 143. Springer, 1979. 5
- [Ren06] J. Renegar. Hyperbolic programs, and their derivative relaxations. Foundations of Computational Mathematics, 6:59–79, 2006. 2
- [San13] R. Sanyal. On the derivative cones of polyhedral cones. Advances in Geometry, 13(2):315–321, 2013. 4
- [Sch23] I. Schur. Uber eine klasse von mittelbildungen mit anwendungen auf die determinantentheorie. Sitzungsberichte der Berliner Mathematischen Gesellschaft, 22:9–20, 1923.
- [SP15] J. Saunderson and P. A. Parrilo. Polynomial-sized semidefinite representations of derivative relaxations of spectrahedral cones. *Mathematical Programming*, 153(2):309–331, 2015. 3, 4, 8
- [Vin12] V. Vinnikov. LMI representations of convex semialgebraic sets and determinantal representations of algebraic hypersurfaces: past, present, and future. In H. Dym, M. C. de Oliveira, and M. Putinar, editors, *Mathematical methods in systems, optimization, and control*, volume 222 of *Operator Theory: Advances and Applications*, pages 325–349. Springer, 2012. 3
- [Zin08] Y. Zinchenko. On hyperbolicity cones associated with elementary symmetric polynomials. *Optimization Letters*, 2(3):389–402, 2008. 4