Polynomial-sized Semidefinite Representations of Derivative Relaxations of Spectrahedral Cones

James Saunderson   Pablo A. Parrilo[*]

August 11, 2014

Abstract

We give explicit polynomial-sized (in n and k) semidefinite representations of the hyperbolicity cones associated with the elementary symmetric polynomials of degree k in n variables. These convex cones form a family of non-polyhedral outer approximations of the non-negative orthant that preserve low-dimensional faces while successively discarding high-dimensional faces. More generally we construct explicit semidefinite representations (polynomial-sized in k, m, and n) of the hyperbolicity cones associated with kth directional derivatives of polynomials of the form $p(x) = \det(\sum_{i=1}^n A_i x_i)$ where the A_i are $m \times m$ symmetric matrices. These convex cones form an analogous family of outer approximations to any spectrahedral cone. Our representations allow us to use semidefinite programming to solve the linear cone programs associated with these convex cones as well as their (less well understood) dual cones.

1 Introduction

Expressing convex optimization problems in conic form, as the minimization of a linear functional over an affine slice of a convex cone, has been an important method in the development of modern convex optimization theory. This abstraction is useful (at least from a theoretical viewpoint) because all that is difficult and interesting about the problem is packaged into the cone. The conic viewpoint provides a natural way to organize classes of convex optimization problems into hierarchies based on whether the cones associated with one class can be expressed in terms of the cones associated with another class. For example, semidefinite programming generalizes linear programming because the non-negative orthant is the restriction to the diagonal of the positive semidefinite cone.

When faced with a convex cone the geometry of which is not well understood, we stand to gain theoretical insight as well as off-the-shelf optimization algorithms by representing it in terms of a cone with known geometric and algebraic structure such as the positive semidefinite cone. Terminology is attached to this idea, with a cone being *spectrahedral* if it is a linear section (or 'slice') of the positive semidefinite cone, and *semidefinitely representable* if it is a linear projection of a spectrahedral cone. The efficiency of a semidefinite representation is also clearly important. If we can write a cone as the projection of a slice of the cone of $m \times m$ positive semidefinite matrices, we say it has a semidefinite representation of *size m*. Many convex cones have been shown to be

---

semidefinitely representable using a variety of techniques (see [14] as well as the recent book [1] for contrasting methods and examples).

The classes of semidefinitely representable cones and spectrahedral cones are distinct [17], with semidefinitely representable cones being perhaps more natural from the point of view of optimization. A semidefinite representation of a cone suffices to express the associated cone program as a semidefinite program. Furthermore, unlike spectrahedral cones, the class of semidefinitely representable cones is closed under duality [4, Proposition 3.2].

The *hyperbolicity cones* form a family of convex cones (constructed from certain multivariate polynomials) that includes the positive semidefinite cone, as well as all homogeneous cones [7]. While it has been shown (by Lewis et al. [10] based on work of Helton and Vinnikov [8]) that all three-dimensional hyperbolicity cones are spectrahedral, little is known about semidefinite representations of higher dimensional hyperbolicity cones. Furthermore while hyperbolicity cones have very simple descriptions, their dual cones are not well understood.

In this paper we give explicit, polynomial-sized semidefinite representations of the hyperbolicity cones known as the *derivative relaxations* of the non-negative orthant, and the corresponding derivative relaxations of the positive semidefinite cone. These cones form a family of outer approximations to the orthant and positive semidefinite cones respectively with many interesting properties [18]. We obtain semidefinite representations of the derivative relaxations of spectrahedral cones as slices of the derivative relaxations of the positive semidefinite cone.

1.1   Hyperbolic polynomials and hyperbolicity cones

A homogeneous polynomial p of degree m in n variables is *hyperbolic* with respect to $e \in \mathbb{R}^n$ if $p(e) \neq 0$ and if for all $x \in \mathbb{R}^n$ the univariate polynomial $t \mapsto p(x - te)$ has only real roots. Gårding's foundational work on hyperbolic polynomials [6] establishes that if p is hyperbolic with respect to e then the connected component of $\{x \in \mathbb{R}^n : p(x) \neq 0\}$ containing e is an open convex cone. This cone is called the *hyperbolicity cone* corresponding to (p, e). We denote it by $\Lambda_{++}(p, e)$, and its closure by $\Lambda_+(p, e)$.

Note that p is hyperbolic with respect to e if and only if $-p$ is hyperbolic with respect to e. As such we assume throughout that $p(e) > 0$. We can expand $p(x + te)$ as

$$p(x + te) = p(e) \left[t^m + a_1(x)t^{m-1} + a_2(x)t^{m-2} + \cdots + a_{m-1}(x)t + a_m(x) \right]$$

where the $a_i(x)$ are polynomials that are homogeneous of degree i. There is an alternative description of the hyperbolicity cone $\Lambda_+(p, e)$ due to Renegar [18, Theorem 20] as

$$\Lambda_+(p, e) = \{x \in \mathbb{R}^n : a_1(x) \geq 0, \ a_2(x) \geq 0, \ \ldots, \ a_m(x) \geq 0\}. \tag{1}$$

We use this description of $\Lambda_+(p, e)$ throughout the paper.

Basic examples:

- The polynomial $p(x_1, x_2, \ldots, x_n) = x_1 x_2 \cdots x_n$ is hyperbolic with respect to $e = \mathbf{1}_n := (1, 1, \ldots, 1)$. The associated closed hyperbolicity cone is the non-negative orthant, $\mathbb{R}^n_+$. Since

$$p(x + t\mathbf{1}_n) = t^n + e_1(x)t^{n-1} + \cdots + e_{n-1}(x)t + e_n(x)$$

  where $e_k(x) = \sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k}$ is the elementary symmetric polynomial of degree k in the variables $x_1, x_2, \ldots, x_n$,

$$\Lambda_+(p, e) = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n : e_1(x) \geq 0, \ e_2(x) \geq 0, \ \ldots, \ e_n(x) \geq 0\}.$$

2

- Let X be an $n \times n$ symmetric matrix of indeterminates. The polynomial $p(X) = \det(X)$ is hyperbolic with respect to $e = I_n$, the $n \times n$ identity matrix. The associated closed hyperbolicity cone is the positive semidefinite cone, $\mathbb{S}_+^n$. Since

$$p(X + tI_n) = t^n + E_1(X)t^{n-1} + \cdots + E_{n-1}(X)t + E_n(X)$$

  where the $E_k(X)$ are the coefficients of the characteristic polynomial of X,

$$\Lambda_+(p, e) = \mathbb{S}_+^n = \{X : E_1(X) \geq 0, \ E_2(X) \geq 0, \ \ldots, \ E_n(X) \geq 0\}.$$

  Observe that $E_k(X) := e_k(\lambda(X))$ is the elementary symmetric polynomial of degree k in the eigenvalues of X so the positive semidefinite cone can also be described in terms of polynomial inequalities on the eigenvalues of X as

$$\mathbb{S}_+^n = \{X : e_1(\lambda(X)) \geq 0, \ e_2(\lambda(X)) \geq 0, \ \ldots, \ e_n(\lambda(X)) \geq 0\}.$$

1.2 Derivative relaxations

If p is hyperbolic with respect to e then (essentially by Rolle's theorem) the directional derivative of p in the direction e, *viz.*

$$p_e^{(1)}(x) := \left.\frac{d}{dt}p(x + te)\right|_{t=0}$$

is also hyperbolic with respect to e, a construction that goes back to Gårding [6]. If p has degree m, by repeatedly differentiating in the direction e we construct a sequence of polynomials $p, p_e^{(1)}, p_e^{(2)}, \ldots, p_e^{(m-1)}$ each hyperbolic with respect to e.

The corresponding hyperbolicity cones can be expressed nicely in terms of polynomial inequalities. Indeed if $p(x + te) = p(e)\left[t^m + \sum_{i=1}^m a_i(x)t^{m-i}\right]$ then differentiating k times with respect to t we see that

$$p_e^{(k)}(x + te) = p(e)\left[c_0 a_{m-k}(x) + c_1 a_{m-k-1}(x)t + \cdots + c_{m-k}t^{m-k}\right]$$

where $c_i = (k + i)!/i! > 0$. By (1) the corresponding hyperbolicity cone is

$$\Lambda_+^{(k)}(p, e) := \Lambda_+(p_e^{(k)}, e) = \{x \in \mathbb{R}^n : a_1(x) \geq 0, \ a_2(x) \geq 0, \ \ldots, \ a_{m-k}(x) \geq 0\}$$

and can be obtained from (1) by removing k of the inequality constraints. As a result, the hyperbolicity cones $\Lambda_+^{(k)}(p, e)$ provide a sequence of outer approximations to the original hyperbolicity cone that satisfy

$$\Lambda_+(p, e) \subset \Lambda_+^{(1)}(p, e) \subset \cdots \subset \Lambda_+^{(m-1)}(p, e).$$

The last of these, $\Lambda_+^{(m-1)}(p, e)$, is simply the closed half-space defined by e. The work of Renegar [18] highlights the many nice properties of this sequence of approximations.

Note that we abuse terminology by referring to the cones $\Lambda_+^{(k)}(p, e)$ as *derivative relaxations* of the hyperbolicity cone $\Lambda_+(p, e)$. The abuse is that $\Lambda_+^{(k)}(p, e)$ does not depend only on the *geometric* object $\Lambda_+(p, e)$ but on its particular *algebraic* description via p and e.

Examples:

- In the case of $p(x) = x_1 x_2 \cdots x_n = e_n(x)$ and $e = \mathbf{1}_n$, we have that $p_e^{(k)}(x) = k! e_{n-k}(x)$. Consequently the kth derivative relaxation of the orthant, which we denote by $\mathbb{R}_+^{n,(k)}$, is the hyperbolicity cone $\Lambda_+(e_{n-k}, \mathbf{1}_n)$. It can be expressed as

$$\mathbb{R}_+^{n,(k)} = \{x \in \mathbb{R}^n : e_1(x) \geq 0, \ e_2(x) \geq 0, \ \ldots, \ e_{n-k}(x) \geq 0\}. \tag{2}$$

  Consistent with these descriptions we define $\mathbb{R}_+^{n,(n)} := \mathbb{R}^n$.

- In the case of $p(X) = \det(X) = E_n(X)$ and $e = I_n$, we have that $p_e^{(k)}(x) = k! E_{n-k}(X)$. The kth derivative relaxation of the positive semidefinite cone, which we denote by $\mathbb{S}_+^{n,(k)}$, can be described as

$$\mathbb{S}_+^{n,(k)} = \{X \in \mathbb{S}^n : E_1(x) \geq 0, \ E_2(x) \geq 0, \ \ldots, \ E_{n-k}(x) \geq 0\} \tag{3}$$

$$= \{X \in \mathbb{S}^n : e_1(\lambda(X)) \geq 0, \ e_2(\lambda(X)) \geq 0, \ \ldots, \ e_{n-k}(\lambda(X)) \geq 0\}. \tag{4}$$

  Again we define $\mathbb{S}_+^{n,(n)} := \mathbb{S}^n$, the set of $n \times n$ symmetric matrices. Since $E_i(\mathrm{diag}(x)) = e_i(x)$ for all i, the diagonal slice of $\mathbb{S}_+^{n,(k)}$ is exactly $\mathbb{R}_+^{n,(k)}$.

Symmetry: Suppose G is a group acting by linear transformations on $\mathbb{R}^n$ by $x \mapsto g \cdot x$ for all $g \in G$. Suppose *both p and e are invariant* under the group action, i.e., $g \cdot e = e$ and $(g \cdot p)(x) := p(g^{-1} \cdot x) = p(x)$ for all $g \in G$. Then for all $t \in \mathbb{R}$, $x \in \mathbb{R}^n$ and $g \in G$

$$p(x + te) = (g \cdot p)(x + te) = p(g^{-1} \cdot (x + te)) = p((g^{-1} \cdot x) + te).$$

Hence the hyperbolicity cone $\Lambda_+(p, e)$ and all of its derivative cones $\Lambda_+^{(k)}(p, e)$ are invariant under this same group action.

For our purposes an important example of this is the symmetry of the cones $\mathbb{S}_+^{n,(k)}$. The action of $O(n)$ by conjugation on symmetric matrices leaves the polynomial $p(X) = \det(X)$ invariant *and* preserves the direction $e = I_n$. Hence all of the derivative relaxations of the positive semidefinite cone are invariant under conjugation by orthogonal matrices. As such, the cones $\mathbb{S}_+^{n,(k)}$ are *spectral sets*, in the sense that whether a symmetric matrix X belongs to $\mathbb{S}_+^{n,(k)}$ depends only on the eigenvalues of X. This is evident from the description of $\mathbb{S}_+^{n,(k)}$ in (4).

1.3 Related work

Previous work has focused on semidefinite and spectrahedral representations of the derivative relaxations of the orthant. Zinchenko [23] used a decomposition approach to give semidefinite representations of $\mathbb{R}_+^{n,(1)}$ and its dual cone. Sanyal [21] subsequently gave spectrahedral representations of $\mathbb{R}_+^{n,(1)}$ and $\mathbb{R}_+^{n,(n-2)}$ and conjectured that all of the derivative relaxations of the orthant admit spectrahedral representations.

Recently Brändén [2] settled this conjecture in the affirmative giving spectrahedral representations of $\mathbb{R}_+^{n,(n-k)}$ for $k = 1, 2, \ldots, n-1$ of size $O(n^{k-1})$. For each $1 \leq k < n$ Brändén constructs a graph $G_{n,k} = (V, E)$ together with edge weights $(w_e(x))_{e \in E}$ that are linear forms in x so that

$$\mathbb{R}_+^{n,(n-k)} = \left\{x \in \mathbb{R}^n : L_{G_{n,k}}(x) \succeq 0\right\} \tag{5}$$

4

where $L_{G_{n,k}}(x)$ is the $|V| \times |V|$ edge-weighted Laplacian of $G_{n,k}$. Since $L_{G_{n,k}}(x)$ is linear in the edge weights, and the edge weights are linear forms in x, (5) is a spectrahedral representation of size $|V|$. With the exception of two distinguished vertices, the vertices of $G_{n,k}$ are indexed by all ℓ-tuples (for $1 \le \ell \le k - 1$) consisting of distinct elements of $\{1, 2, \ldots, n\}$. Hence $|V| = 2 + \sum_{\ell=1}^{k-1} \ell! \binom{n}{\ell}$ showing that Brändén's spectrahedral representation of $\mathbb{R}_+^{n,(n-k)}$ has size $O(n^{k-1})$. While Brändén's construction is of considerable theoretical interest, these representations (unlike ours) are not practical for optimization due to their prohibitive size.

A spectrahedral representation of $\mathbb{R}_+^{n,(1)}$ is implicit in the work of Choe et al. [3] that studies the relationships between matroids and hyperbolic polynomials. Choe et al. observe that if $\mathcal{M}$ is a *regular matroid* represented by the rows of a totally unimodular matrix V then $\det(V^T \operatorname{diag}(x) V)$ is the basis generating polynomial of $\mathcal{M}$. In particular, the uniform matroid U_n^{n-1} is regular and has $e_{n-1}(x)$ as its basis generating polynomial, yielding a symmetric determinantal representation of $e_{n-1}(x)$ and hence a spectrahedral representation of $\mathbb{R}_+^{n,(n-1)}$.

From a computational perspective, Güler [7] showed that if p has degree m and is hyperbolic with respect to e then $\log p$ is a self-concordant barrier function (with barrier parameter m) for the hyperbolicity cone $\Lambda_+(p, e)$. As such, as long as p and its gradient and Hessian can be computed efficiently, one can use interior point methods to minimize a linear functional over an affine slice of $\Lambda_+(p, e)$ efficiently. Renegar [18, Section 9] gave an efficient interpolation-based method for computing $p_e^{(k)}$ (and its gradient and Hessian) whenever p (and its gradient and Hessian) can be evaluated efficiently. Güler and Renegar's observations together yield efficient computational methods to optimize a linear functional over an affine slice of a derivative relaxation of a spectrahedral cone. Our results complement these, giving a method to solve optimization problems of this type using existing numerical procedures for semidefinite programming.

1.4 Notation

Here we define notation not explicitly defined elsewhere in the paper. If C is a convex cone, we denote by C^* the dual cone, i.e. the set of linear functionals that are non-negative on C. We represent linear functionals on $\mathbb{R}^n$ using the standard Euclidean inner product, and linear functionals on $\mathbb{S}^n$ using the trace inner product $\langle X, Y \rangle = \operatorname{tr}(XY)$. As such $C^* = \{y : \langle y, x \rangle \ge 0, \text{ for all } x \in C\}$. If $X \in \mathbb{S}^n$ let $\lambda(X)$ denote its eigenvalues sorted so that $\lambda_1(X) \ge \lambda_2(X) \ge \cdots \ge \lambda_n(X)$. If $X \in \mathbb{S}^n$ let $\operatorname{diag}(X) \in \mathbb{R}^n$ denote the vector of diagonal entries and if $x \in \mathbb{R}^n$ let $\operatorname{diag}(x)$ denote the diagonal matrix with diagonal entries given by x. The usage will be clear from the context.

2 Results

Our main contribution is to construct two different explicit polynomial-sized semidefinite representations of the derivative relaxations of the positive semidefinite cone. We call our two representations the *derivative-based* and *polar derivative-based* representations respectively. In this section we describe these representations, and outline the proof of our main theoretical result.

Theorem 1. *For each positive integer n and each $k = 1, 2, \ldots, n-1$, the cone $\mathbb{S}_+^{n,(k)}$ has a semidefinite representation of size $O(\min\{k, n-k\}n^2)$.*

We defer detailed proofs of the correctness of our representations to Sections 3 and 4. At this stage, we just highlight that there is essentially one basic algebraic fact that underlies all of our

results. Whenever V_n is an $n \times (n-1)$ matrix with orthonormal columns that are each orthogonal to $\mathbf{1}_n$, i.e. $V_n^T V_n = I_{n-1}$ and $V_n^T \mathbf{1}_n = 0$, then

$$e_{n-1}(x) = n \det(V_n^T \operatorname{diag}(x) V_n).$$

We give a proof of this identity in Section 3. Note that this identity is independent of the particular choice of V_n satisfying $V_n^T V_n = I_{n-1}$ and $V_n^T \mathbf{1}_n = 0$. In fact, all of the results expressed in terms of V_n (notably Propositions 2, 3, 2D, and 3D) are similarly independent of the particular choice of V_n.

Both of the representations are recursive in nature. The derivative-based representation is based on recursively applying two basic propositions (Propositions 1 and 2, to follow) to construct a chain of semidefinite representations of the form

$$\boxed{\mathbb{S}_+^{n,(k)}} \xleftarrow[\text{Prop. 1}]{O(n^2)} \mathbb{R}_+^{n,(k)} \xleftarrow[\text{Prop. 2}]{0} \boxed{\mathbb{S}_+^{n-1,(k-1)}} \xleftarrow[\text{Prop. 1}]{O((n-1)^2)} \mathbb{R}_+^{n-1,(k-1)} \leftarrow \cdots \tag{6}$$

$$\cdots \leftarrow \mathbb{R}_+^{n-k+1,(1)} \xleftarrow[\text{Prop. 2}]{0} \boxed{\mathbb{S}_+^{n-k,(0)}}.$$

The annotated arrow $C \xleftarrow[\text{Prop. } a]{m} K$ indicates that given a semidefinite representation of K of size m' we can construct a semidefinite representation of C of size $m' + m$, and that an explicit description of the construction is given in Proposition a.

The base case of the recursion is just the positive semidefinite cone $\mathbb{S}_+^{n-k,(0)}$, which has a trivial semidefinite representation. Hence starting from $\mathbb{S}_+^{n-k,(0)}$ (which has a semidefinite representation of size $n - k$), we can apply Proposition 2 to obtain a semidefinite representation of $\mathbb{R}_+^{n-k+1,(1)}$ of size $n - k$, then apply Proposition 1 to obtain a semidefinite representation of $\mathbb{S}_+^{n-k+1,(1)}$ of size $(n - k) + O((n - k + 1)^2)$, and so on.

The polar derivative-based representation is based on recursively applying Proposition 1 together with a third basic proposition (Proposition 3, to follow) to construct a slightly different chain of semidefinite representations of the form

$$\boxed{\mathbb{S}_+^{n,(k)}} \xleftarrow[\text{Prop. 1}]{O(n^2)} \mathbb{R}_+^{n,(k)} \xleftarrow[\text{Prop. 3}]{n} \boxed{\mathbb{S}_+^{n-1,(k)}} \xleftarrow[\text{Prop. 1}]{O(n^2)} \mathbb{R}_+^{n-1,(k)} \leftarrow \cdots$$

$$\cdots \leftarrow \mathbb{R}_+^{k+2,(k)} \xleftarrow[\text{Prop. 3}]{n} \boxed{\mathbb{S}_+^{k+1,(k)}}. \tag{7}$$

Note that the base case of the recursion is just $\mathbb{S}_+^{k+1,(k)} = \{X \in \mathbb{S}^{k+1} : \operatorname{tr}(X) \geq 0\}$, a half-space.

2.1 Building blocks of the two recursions

We now describe the constructions related to each of the types of arrows in the recursions sketched above. The arrows labeled by Proposition 1 assert that we can construct a semidefinite representation of $\mathbb{S}_+^{n,(k)}$ from a semidefinite representation of $\mathbb{R}_+^{n,(k)}$. This can be done in the following way.

Proposition 1. *If $\mathbb{R}_+^{n,(k)}$ has a semidefinite representation of size m, then $\mathbb{S}_+^{n,(k)}$ has a semidefinite representation of size $m + O(n^2)$. Indeed*

$$\mathbb{S}_+^{n,(k)} = \left\{ X \in \mathbb{S}^n : \exists z \in \mathbb{R}^n \ \text{s.t.} \ z \in \mathbb{R}_+^{n,(k)}, \ (X, z) \in \operatorname{SH}_n \right\}, \tag{8}$$

where SH_n *is the* Schur-Horn cone *defined as*

$$\mathrm{SH}_n = \big\{(X, z) : \ z_1 \geq z_2 \geq \cdots \geq z_n, \ \ X \in \mathrm{conv}_{Q \in O(n)} \{Q^T \operatorname{diag}(z)Q\}\big\}$$

i.e. the set of pairs (X, z) *such that* X *is in the convex hull of all symmetric matrices with ordered spectrum* z. *The Schur-Horn cone has the semidefinite characterization*

$$
\begin{aligned}
(X, z) \in \mathrm{SH}_n \quad &\textit{if and only if} \quad z_1 \geq z_2 \geq \cdots \geq z_n \ \ \textit{and} \\
&\textit{there exist} \quad t_2, \ldots, t_{n-1} \in \mathbb{R}, \ \ Z_2, \ldots, Z_{n-1} \succeq 0 \\
&\textit{such that} \quad \operatorname{tr}(X) = \textstyle\sum_{j=1}^n z_j, \ \ X \preceq z_1 I, \ \ \textit{and} \\
&\textit{for } \ell = 2, \ldots, n-1, \quad X \preceq t_\ell I + Z_\ell \ \ \textit{and} \ \ \ell \cdot t_\ell + \operatorname{tr}(Z_\ell) \leq \textstyle\sum_{j=1}^\ell z_j.
\end{aligned}
$$

Proposition 1 holds because of the *symmetry* of $\mathbb{S}_+^{n,(k)}$. In particular it is a spectral set—invariant under conjugation by orthogonal matrices. The other reason this representation works is that the diagonal slice of $\mathbb{S}_+^{n,(k)}$ is $\mathbb{R}_+^{n,(k)}$. We discuss this result in more detail in Section 4.

The arrows in (6) labeled by Proposition 2 appear only in the derivative-based recursion. They assert that we can obtain a semidefinite representation of $\mathbb{R}_+^{n,(k)}$ from a semidefinite representation of $\mathbb{S}_+^{n-1,(k-1)}$. Indeed we establish in Section 3.1 that $\mathbb{R}_+^{n,(k)}$ is actually a slice of $\mathbb{S}_+^{n-1,(k-1)}$.

Proposition 2. *If* $1 \leq k \leq n-1$ *then* $\mathbb{R}_+^{n,(k)} = \left\{ x \in \mathbb{R}^n : V_n^T \operatorname{diag}(x) V_n \in \mathbb{S}_+^{n-1,(k-1)} \right\}$.

The arrows in (7) labeled by Proposition 3 appear only in the polar derivative-based recursion. They assert that we can obtain a semidefinite representation of $\mathbb{R}_+^{n,(k)}$ from a semidefinite representation of $\mathbb{S}_+^{n-1,(k)}$. We establish the following in Section 3.2.

Proposition 3. *If* $1 \leq k \leq n-2$ *then*

$$\mathbb{R}_+^{n,(k)} = \left\{ x \in \mathbb{R}^n : \exists Z \in \mathbb{S}_+^{n-1,(k)} \ \ s.t. \ \ \operatorname{diag}(x) \succeq V_n Z V_n^T \right\}.$$

2.2 Size of the representations

Recall that each arrow $C \xleftarrow{m} K$ in (6) and (7) is labeled with the *additional size* m required to implement the representation of C given a semidefinite representation of K. Since the derivative-based recursion has $2k$ arrows, it is immediate from (6) that the derivative-based semidefinite representation of $\mathbb{S}_+^{n,(k)}$ has size $O(kn^2)$ and so is of polynomial size.

On the other hand, this approach gives a disappointingly large semidefinite representation of the half-space $\mathbb{S}_+^{n,(n-1)} = \{X \in \mathbb{S}^n : \operatorname{tr}(X) \geq 0\}$ of size $O(n^3)$. The derivative-based approach cannot exploit the fact that this is a very simple cone. This is why we also consider the polar derivative-based representation, as it is designed around the fact that $\mathbb{S}_+^{n,(n-1)}$ has a simple semidefinite representation.

It is immediate from (7) that the polar derivative-based semidefinite representation of $\mathbb{S}_+^{n,(k)}$ has size $O((n-k)n^2)$ and so is also of polynomial size. Furthermore, it gives small representations of size $O(n^2)$ exactly when the derivative-based representations are large, of size $O(n^3)$. For any given pair (n, k) we should always use the derivative-based representation of $\mathbb{S}_+^{n,(k)}$ if $k < n/2$ and the polar derivative-based representation when $k > n/2$. Theorem 1 combines our two size estimates, stating that $\mathbb{S}_+^{n,(k)}$ has a semidefinite representation of size $O(\min\{k, n-k\}n^2)$.

2.3 Pseudocode for our derivative-based representation

We do not write out any of our semidefinite representations in full because the recursive descriptions given here are actually more naturally suited to implementation. To illustrate this, we give pseudocode for the MATLAB-based high-level modeling language YALMIP [11] that 'implements' the derivative-based representations of $\mathbb{S}_+^{n,(k)}$ and $\mathbb{R}_+^{n,(k)}$. Decision variables are declared by expressions like x = sdpvar(n,1); which creates a decision variable x taking values in $\mathbb{R}^n$. An LMI object is a list of equality constraints and linear matrix inequality constraints that are linear in any declared decision variables.

Suppose we have a function SH(X,z) that takes a pair of decision variables and returns an LMI object corresponding to the constraint that $(X, z) \in \mathrm{SH}_n$. This is easy to construct from the explicit semidefinite representation in Proposition 1. Then the function psdcone takes an $n \times n$ symmetric matrix-valued decision variable X and returns an LMI object for the constraint $X \in \mathbb{S}_+^{n,(k)}$.

```
1:  function K = psdcone(X,k)
2:      if k==0
3:          K = [X >= 0];
4:      else
5:          z = sdpvar(size(X,1),1);
6:          K = [orthant(z,k), SH(X,z)];
7:      end
```

It calls a function orthant that takes a decision variable x in $\mathbb{R}^n$ and returns an LMI object for the constraint $x \in \mathbb{R}_+^{n,(k)}$.

```
1:  function K = orthant(x,k)
2:      if k==0
3:          K = [x >= 0];
4:      else
5:          V = null(ones(size(x))');
6:          K = [psdcone(V'*diag(x)*V,k-1)];
7:      end
```

It is straightforward to adapt these two functions for the polar derivative-based representation, one needs only to change the base cases (lines 2–4 of each) and to adapt line 6 of orthant to reflect Proposition 3.

2.4 Dual cones

If a cone is semidefinitely representable, so is its dual cone. In fact there are explicit procedures to take a semidefinite representation for a cone and produce a semidefinite representation for its dual cone [13, Section 4.1.1]. Here we describe two explicit semidefinite representations of the dual cones $(\mathbb{S}_+^{n,(k)})^*$ that enjoy the same recursive structure as the corresponding semidefinite representations of $\mathbb{S}_+^{n,(k)}$.

To construct them, we essentially dualize all the relationships given by the arrows in (6) and (7). By straightforward applications of a conic duality argument, in Section 3.3 we establish the following dual analogues of Propositions 2 and 3.

Proposition 2D. If $1 \leq k \leq n-1$ then

$$(\mathbb{R}_+^{n,(k)})^* = \left\{ \mathrm{diag}(V_n Y V_n^T) : \ Y \in (\mathbb{S}_+^{n-1,(k-1)})^* \right\}.$$

Proposition 3D. If $1 \leq k \leq n-2$ then

$$(\mathbb{R}_+^{n,(k)})^* = \left\{ \mathrm{diag}(Y) : \ Y \succeq 0, \ V_n^T Y V_n \in (\mathbb{S}_+^{n-1,(k)})^* \right\}.$$

We could also obtain a dual version of Proposition 1 by directly applying conic duality to the semidefinite representation in Proposition 1. This would involve dualizing the semidefinite representation of SH_n. Instead we give another, perhaps simpler, representation of $(\mathbb{S}_+^{n,(k)})^*$ in terms of $(\mathbb{R}_+^{n,(k)})^*$ that is not obtained by directly applying conic duality to Proposition 1.

Proposition 1D. If $(\mathbb{R}_+^{n,(k)})^*$ has a semidefinite representation of size m, then $(\mathbb{S}_+^{n,(k)})^*$ has a semidefinite representation of size $m + O(n^2)$ given by

$$(\mathbb{S}_+^{n,(k)})^* = \left\{ W \in \mathbb{S}^n : \exists y \in \mathbb{R}^n \ \mathrm{s.t.} \ y \in (\mathbb{R}_+^{n,(k)})^*, \ (W,y) \in \mathrm{SH}_n \right\}. \tag{9}$$

Recall that Proposition 1 holds because $\mathbb{S}_+^{n,(k)}$ is invariant under orthogonal conjugation and $\mathbb{R}_+^{n,(k)}$ is the diagonal slice of $\mathbb{S}_+^{n,(k)}$. While it is immediate that $(\mathbb{S}_+^{n,(k)})^*$ is also orthogonally invariant, it is a less obvious result that the diagonal slice of $(\mathbb{S}_+^{n,(k)})^*$ is $(\mathbb{R}_+^{n,(k)})^*$. We prove this in Section 4.

The recursions underlying the derivative-based and polar derivative-based representations of $(\mathbb{S}_+^{n,(k)})^*$ then take the form

$$(\mathbb{S}_+^{n,(k)})^* \leftarrow (\mathbb{R}_+^{n,(k)})^* \leftarrow (\mathbb{S}_+^{n-1,(k-1)})^* \leftarrow \cdots \leftarrow (\mathbb{R}_+^{n-k+1,(1)})^* \leftarrow (\mathbb{S}_+^{n-k,(0)})^* \tag{10}$$

and, respectively,

$$(\mathbb{S}_+^{n,(k)})^* \leftarrow (\mathbb{R}_+^{n,(k)})^* \leftarrow (\mathbb{S}_+^{n-1,(k)})^* \leftarrow \cdots \leftarrow (\mathbb{R}_+^{k+2,(k)})^* \leftarrow (\mathbb{S}_+^{k+1,(k)})^*. \tag{11}$$

Note that for the dual derivative-based representation, the base case is $(\mathbb{S}_+^{n-k,(0)})^* = \mathbb{S}_+^{n-k}$ (since the positive semidefinite cone is self dual). For the dual polar derivative-based representation the base case is $(\mathbb{S}_+^{k+1,(k)})^* = \{t I_{k+1} : t \geq 0\}$, the ray generated by the identity matrix in $\mathbb{S}^{k+1}$.

2.5 Derivative relaxations of spectrahedral cones

So far we have focused on the derivative relaxations of the positive semidefinite cone. It turns out that the derivative relaxations of spectrahedral cones are just slices of the associated derivative relaxations of the positive semidefinite cone.

Proposition 4. *Suppose* $p(x) = \det(\sum_{i=1}^n A_i x_i)$ *where the* A_i *are* $m \times m$ *symmetric matrices and* $e \in \mathbb{R}^n$ *is such that* $\sum_{i=1}^n A_i e_i = B$ *is positive definite. Then for* $k = 0, 1, \ldots, m-1$,

$$\Lambda_+^{(k)}(p,e) = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n B^{-1/2} A_i B^{-1/2} x_i \in \mathbb{S}_+^{m,(k)} \right\}.$$

9

Proof. Let $A(x) = \sum_{i=1}^{n} B^{-1/2} A_i B^{-1/2} x_i$. Then $A(e) = I$ and for all $x \in \mathbb{R}^n$ and all $t \in \mathbb{R}$

$$p(x + te) = \det(B) \det(A(x + te)) = \det(B) \det(A(x) + tI).$$

This implies that all the derivatives of p in the direction e are exactly the same as the corresponding derivatives of $\det(B) \det(X)$ in the direction I evaluated at $X = A(x)$. Since $\det(B) > 0$, it follows that for $k = 0, 1, \ldots, m - 1$, $x \in \Lambda_+^{(k)}(p, e)$ if and only if $A(x) \in \mathbb{S}_+^{m,(k)}$. $\qquad\square$

We conclude this section with an example of these constructions.

Example 1 (Derivative relaxations of a 3-ellipse)**.** Given foci $(0, 0), (0, 4)$ and $(3, 0)$ in the plane, the 3-ellipse consisting of points such that the sum of distances to the foci equals 8 is shown in Figure 1. This is one connected component of the real algebraic curve of degree 8 given by $\{(x, y) \in \mathbb{R}^2 : \det \mathcal{E}(x, y, 1) = 0\}$ where $\mathcal{E}$ is defined in (12) (see Nie et al. [16]). The region enclosed by this 3-ellipse is the $z = 1$ slice of the spectrahedral cone defined by $\mathcal{E}(x, y, z) \succeq 0$ where

$$\mathcal{E}(x, y, z) = \begin{bmatrix} 5z + 3x & y & y - 4z & 0 & y & 0 & 0 & 0 \\ y & 5z + x & 0 & y - 4z & 0 & y & 0 & 0 \\ y - 4z & 0 & 5z + x & y & 0 & 0 & y & 0 \\ 0 & y - 4z & y & 5z - x & 0 & 0 & 0 & y \\ y & 0 & 0 & 0 & 11z + x & y & y - 4z & 0 \\ 0 & y & 0 & 0 & y & 11z - x & 0 & y - 4z \\ 0 & 0 & y & 0 & y - 4z & 0 & 11z - x & y \\ 0 & 0 & 0 & y & 0 & y - 4z & y & 11z - 3x \end{bmatrix}. \quad (12)$$

Note that $\mathcal{E}(0, 0, 1) \succ 0$ and so $e = (0, 0, 1)$ is a direction of hyperbolicity for $p(x, y, z) = \det \mathcal{E}(x, y, z)$. The left of Figure 1 shows the $z = 1$ slice of the cone $\Lambda_+(p, e)$ and its first three derivative relaxations $\Lambda_+^{(1)}(p, e), \Lambda_+^{(2)}(p, e)$, and $\Lambda_+^{(3)}(p, e)$. The right of Figure 1 shows the $z = 1$ slice of the cones $(\Lambda_+(p, e))^*, (\Lambda_+^{(1)}(p, e))^*, (\Lambda_+^{(2)}(p, e))^*$, and $(\Lambda_+^{(3)}(p, e))^*$. All of these convex bodies were plotted by computing 200 points on their respective boundaries by optimizing 200 different linear functionals over them. We performed the optimization by modeling our semidefinite representations of these cones in YALMIP [11] which numerically solved the corresponding semidefinite program using SDPT3 [22].

3 The derivative-based and polar derivative-based recursive constructions

In this section we prove Proposition 2 which relates $\mathbb{R}_+^{n,(k)}$ and $\mathbb{S}_+^{n-1,(k-1)}$ as well as Proposition 3 which relates $\mathbb{R}_+^{n,(k)}$ and $\mathbb{S}_+^{n-1,(k)}$. These relationships are the geometric consequences of polynomial identities between elementary symmetric polynomials and determinants.

Specifically the proof of Proposition 2 makes use of a determinantal representation (Equation (15) in Section 3.1) of the derivative

$$\tfrac{\partial}{\partial t} e_n(sx + t\mathbf{1}_n)\big|_{s=1} = \left[1 \cdot e_{n-1}(x) + \cdots + (n-1) \cdot e_1(x) t^{n-2} + n \cdot t^{n-1} \right]. \quad (13)$$

(Note that s plays no role in (13), we include it to highlight the relationship with (14).) Similarly the proof of Proposition 3 relies on a determinantal expression (Equation (18) in Section 3.2) for the polar derivative

$$\tfrac{\partial}{\partial s} e_n(sx + t\mathbf{1}_n)\big|_{s=1} = \left[n \cdot e_n(x) + (n-1) \cdot e_{n-1}(x) t + \cdots + 1 \cdot e_1(x) t^{n-1} \right]. \quad (14)$$
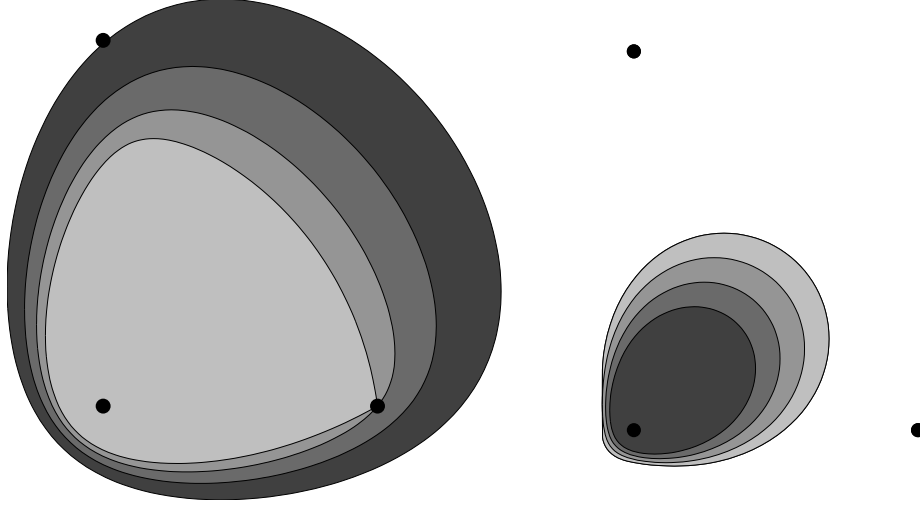
Figure 1: On the left, the inner region is the 3-ellipse consisting of points with sum-of-distances to $(0,0), (0,4),$ and $(3,0)$ equal to 8, i.e. the $z = 1$ slice of the spectrahedral cone defined by (12). The outer three regions are the $z = 1$ slices of the first three derivative relaxations of this spectrahedral cone in the direction $(0,0,1)$. On the right are the $z = 1$ slices of the dual cones of the cones shown on the left, with dual pairs having the same shading.

This explains why we call one the *derivative-based representation*, and the other the *polar derivative-based representation*.

3.1 The derivative-based recursion: relating $\mathbb{R}_+^{n,(k)}$ and $\mathbb{S}_+^{n-1,(k-1)}$

Let V_n denote an (arbitrary) $n \times (n-1)$ matrix satisfying $V_n^T V_n = I_{n-1}$ and $V_n^T \mathbf{1}_n = 0$. Our results in this section and the next stem from the following identity.

Lemma 1. *For all $x \in \mathbb{R}^n$ and all $t \in \mathbb{R}$,*

$$\tfrac{\partial}{\partial t} e_n(sx + t\mathbf{1}_n)\big|_{s=1} = e_{n-1}(x + t\mathbf{1}_n) = n \det(V_n^T \operatorname{diag}(x) V_n + tI_{n-1}). \tag{15}$$

This is a special case of an identity established by Choe et al. [3, Corollary 8.2] and is closely related to Sanyal's result [21, Theorem 1.1]. The proof of Choe et al. uses the Cauchy-Binet identity. Here we provide an alternative proof.

Proof. The polynomial $e_{n-1}(x_1, x_2, \ldots, x_n)$ is characterized by satisfying $e_{n-1}(\mathbf{1}_n) = n$, and by being symmetric, homogeneous of degree $n-1$ and of degree one in each of the x_i. We show, below, that $n \det(V_n^T \operatorname{diag}(x) V_n)$ also has these properties and so that $e_{n-1}(x) = n \det(V_n^T \operatorname{diag}(x) V_n)$. The stated result then follows because $V_n^T V_n = I_{n-1}$ implies

$$e_{n-1}(x + t\mathbf{1}_n) = n \det(V_n^T \operatorname{diag}(x + t\mathbf{1}_n) V_n) = n \det(V_n^T \operatorname{diag}(x) V_n + tI_{n-1}).$$

Now, it is clear that $\det(V_n^T \operatorname{diag}(x) V_n)$ is homogeneous of degree $n-1$ and that

$$n \det(V_n^T \operatorname{diag}(\mathbf{1}_n) V_n) = n \det(I_{n-1}) = n.$$

11

It remains to establish that $\det(V_n^T \operatorname{diag}(x)V_n)$ is symmetric and of degree one in each of the x_i. To do so we repeatedly use the fact that if V_n and U_n both have orthonormal columns that span the orthogonal complement of $\mathbf{1}_n$ then $\det(V_n^T \operatorname{diag}(x)V_n) = \det(U_n^T \operatorname{diag}(x)U_n)$.

The polynomial $\det(V_n^T \operatorname{diag}(x)V_n)$ is symmetric because for any $n \times n$ permutation matrix P the columns of V_n and PV_n respectively are both orthonormal and each spans the orthogonal complement of $\mathbf{1}_n$ (because $P\mathbf{1}_n = \mathbf{1}_n$). Hence

$$\det(V_n^T \operatorname{diag}(Px)V_n) = \det((PV_n)^T \operatorname{diag}(x)(PV_n)) = \det(V_n^T \operatorname{diag}(x)V_n).$$

We finally show that $\det(V_n^T \operatorname{diag}(x)V_n)$ is of degree one in each x_i by a convenient choice of V_n. For any i, we can always choose V_n to be of the form

$$V_n^T = \begin{bmatrix} v_1 & \cdots v_{i-1} & \sqrt{\frac{n-1}{n}}e_i & v_{i+1} & \cdots & v_n \end{bmatrix}$$

where e_i is the ith standard basis vector in $\mathbb{R}^{n-1}$. Then

$$\det(V_n^T \operatorname{diag}(x)V_n) = \det\left(x_i \left(\tfrac{n-1}{n}\right) e_i e_i^T + \sum_{j \neq i} x_j v_j v_j^T \right)$$

which is of degree one in x_i by the linearity of the determinant in its ith column. □

As observed by Sanyal, such a determinantal identity for $e_{n-1}(x)$ establishes that $\mathbb{R}_+^{n,(1)}$ is a slice of $\mathbb{S}_+^{n-1} = \mathbb{S}_+^{n-1,(1-1)}$. We now have two expressions for the derivative $\frac{\partial}{\partial t}e_n(sx + t\mathbf{1}_n)\big|_{s=1}$, one from the definition (13) and one from (15). Comparing them allows us to deduce Proposition 2, that $\mathbb{R}_+^{n,(k)}$ is a slice of $\mathbb{S}_+^{n-1,(k-1)}$ for all $1 \leq k \leq n-1$.

of Proposition 2. From (13) and (15) we see that

$$\frac{\partial}{\partial t}e_n(sx + t\mathbf{1}_n)\big|_{s=1} = \left[1 \cdot e_{n-1}(x) + \cdots + (n-1) \cdot e_1(x)t^{n-2} + n \cdot t^{n-1} \right]$$
$$= n \left[E_{n-1}(V_n^T \operatorname{diag}(x)V_n) + \cdots + E_1(V_n^T \operatorname{diag}(x)V_n)t^{n-2} + t^{n-1} \right].$$

Comparing coefficients of powers of t we see that for $i = 0, 1, \ldots, n-1$

$$nE_{(n-1)-(i-1)}(V_n^T \operatorname{diag}(x)V_n) = (n-i)e_{n-i}(x).$$

Hence for $k = 1, 2, \ldots, n-1$, $x \in \mathbb{R}_+^{n,(k)}$ if and only if $V_n^T \operatorname{diag}(x)V_n \in \mathbb{S}_+^{n-1,(k-1)}$. □

3.2 The polar derivative-based recursion: relating $\mathbb{R}_+^{n,(k)}$ and $\mathbb{S}_+^{n-1,(k)}$

In this section we relate $\mathbb{R}_+^{n,(k)}$ with $\mathbb{S}_+^{n-1,(k)}$, eventually proving Proposition 3. Our argument follows a pattern similar to the previous section. First we give a determinantal expression for the polar derivative $\frac{\partial}{\partial s}e_n(sx + t\mathbf{1}_n)\big|_{s=1}$, and then interpret it geometrically.

While our approach here is closely related to the approach of the previous section, things are a little more complicated. This is not surprising because our construction aims to express $\mathbb{R}_+^{n,(k)}$, which has an algebraic boundary of degree $n-k$, in terms of $\mathbb{S}_+^{n-1,(k)}$, which has an algebraic boundary of *smaller* degree, $n-k-1$. Hence it is not possible for $\mathbb{R}_+^{n,(k)}$ simply to be a slice of $\mathbb{S}_+^{n-1,(k)}$.

Block matrix notation: Let $\hat{\mathbf{1}}_n = \mathbf{1}_n/\sqrt{n}$ and define $Q_n = \begin{bmatrix} V_n & \hat{\mathbf{1}}_n \end{bmatrix}$ noting that Q_n is orthogonal. It is convenient to introduce the block matrix

$$M(x) := Q_n^T \operatorname{diag}(x) Q_n = \begin{bmatrix} V_n^T \operatorname{diag}(x) V_n & V_n^T \operatorname{diag}(x) \hat{\mathbf{1}}_n \\ \hat{\mathbf{1}}_n^T \operatorname{diag}(x) V_n & \hat{\mathbf{1}}_n^T \operatorname{diag}(x) \hat{\mathbf{1}}_n \end{bmatrix} =: \begin{bmatrix} M_{11}(x) & M_{12}(x) \\ M_{12}(x)^T & M_{22}(x) \end{bmatrix} \tag{16}$$

which reflects the fact that it is natural to work in coordinates that are adapted to the symmetry of the problem. (Indeed $\hat{\mathbf{1}}_n$ and the columns of V_n each span invariant subspaces for the permutation action on the coordinates of $\mathbb{R}^n$.)

Schur complements: In this section our results are expressed naturally in term of the *Schur complement* $(M/M_{22})(x) := M_{11}(x) - M_{12}(x) M_{22}(x)^{-1} M_{12}(x)^T$ which is well defined whenever $e_1(x) = n M_{22}(x) \neq 0$. The following lemma summarizes the main properties of the Schur complement that we use.

Lemma 2. If $M = \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix}$ is a partitioned symmetric matrix with non-zero scalar M_{22} and $M/M_{22} := M_{11} - M_{12} M_{22}^{-1} M_{12}^T$ then

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix} = \begin{bmatrix} I_{n-1} & M_{12} M_{22}^{-1} \\ 0 & I_1 \end{bmatrix} \begin{bmatrix} M/M_{22} & 0 \\ 0 & M_{22} \end{bmatrix} \begin{bmatrix} I_{n-1} & 0 \\ M_{22}^{-1} M_{12}^T & I_1 \end{bmatrix}. \tag{17}$$

This factorization immediately implies the following properties.

- *If M is invertible then the $(1,1)$ block of M^{-1} is given by $[M^{-1}]_{11} = (M/M_{22})^{-1}$.*

- *If $M_{22} > 0$ then*

$$M \succeq 0 \iff M/M_{22} \succeq 0.$$

We now establish our determinantal expression for the polar derivative.

Lemma 3. If $e_1(x) = n M_{22}(x) \neq 0$ then

$$\left.\tfrac{\partial}{\partial s} e_n(sx + t\mathbf{1}_n)\right|_{s=1} = e_1(x) \det((M/M_{22})(x) + t I_{n-1}). \tag{18}$$

Proof. First assume $x_i \neq 0$ for $i = 1, 2, \ldots, n$. If $x \in \mathbb{R}^n$ let x^{-1} denote its entry-wise inverse. Exploiting our determinantal expression for the derivative we see that

$$\begin{aligned}
\tfrac{\partial}{\partial s} e_n(sx + t\mathbf{1}_n) &= e_n(x) \tfrac{\partial}{\partial s} e_n(s\mathbf{1}_n + tx^{-1}) \\
&\overset{*}{=} e_n(x)\, n\, \det(V_n^T \operatorname{diag}(tx^{-1} + s\mathbf{1}_n) V_n) \\
&= e_n(x)\, n\, \det(V_n^T \operatorname{diag}(x^{-1}) V_n) \det(t I_{n-1} + s(V_n^T \operatorname{diag}(x^{-1}) V_n)^{-1}) \\
&\overset{*}{=} e_n(x) e_{n-1}(x^{-1}) \det(t I_{n-1} + s(V_n^T \operatorname{diag}(x^{-1}) V_n)^{-1}) \\
&= e_1(x) \det(t I_{n-1} + s(V_n^T \operatorname{diag}(x^{-1}) V_n)^{-1}) \tag{19}
\end{aligned}$$

where the equalities marked with an asterisk are due to (15). Since Q_n is orthogonal $M(x)^{-1} = (Q_n^T \operatorname{diag}(x) Q_n)^{-1} = Q_n^T \operatorname{diag}(x^{-1}) Q_n = M(x^{-1})$. Hence using a property of the Schur complement from Lemma 2 we see that

$$(V_n^T \operatorname{diag}(x^{-1}) V_n)^{-1} = [M(x^{-1})]_{11}^{-1} = [M(x)^{-1}]_{11}^{-1} = (M/M_{22})(x).$$

Substituting this into (19) establishes the stated identity, which, by continuity, is valid for all x such that $e_1(x) = n M_{22}(x) \neq 0$. $\qquad \square$

We now have two expressions for the polar derivative, namely (14) and (18). One comes from the definition of polar derivative, the other from the determinantal representation of Lemma 3. Expanding each and equating coefficients gives the following identities.

Lemma 4. *Let $x \in \mathbb{R}^n$ be such that $e_1(x) = nM_{22}(x) \neq 0$. Then for $k = 0, 1, 2, \ldots, n-1$*

$$e_1(x)E_{n-1-k}((M/M_{22})(x)) = (n-k)e_{n-k}(x).$$

Proof. Expanding the polar derivative two ways (from Lemma 3 and (14)) we obtain

$$\frac{\partial}{\partial s}e_n(sx + t\mathbf{1}_n)\big|_{s=1} = \left[n \cdot e_n(x) + (n-1) \cdot e_{n-1}(x)t + \cdots + 1 \cdot e_1(x)t^{n-1}\right]$$
$$= e_1(x)\left[E_{n-1}((M/M_{22})(x)) + E_{n-2}((M/M_{22})(x))t + \cdots + t^{n-1}\right].$$

The result follows by equating coefficients of t^k. □

We are now in a position to prove the main result of this section.

of Proposition 3. From the definition of $M(x)$ in (16), observe that because Q_n is orthogonal, the constraint $\operatorname{diag}(x) \succeq V_n Z V_n^T$ holds if and only if

$$M(x) = Q_n^T \operatorname{diag}(x)Q_n \succeq Q_n^T(V_n Z V_n^T)Q_n = \begin{bmatrix} Z & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence we aim to establish the following statement that is equivalent to Proposition 3

$$\mathbb{R}_+^{n,(k)} = \left\{ x \in \mathbb{R}^n : \exists Z \in \mathbb{S}_+^{n-1,(k)} \text{ s.t. } M(x) \succeq \begin{bmatrix} Z & 0 \\ 0 & 0 \end{bmatrix} \right\} \quad \text{for } k = 1, 2, \ldots, n-2.$$

The arguments that follow repeatedly use the fact (from Lemma 2) that if $e_1(x) = nM_{22}(x) > 0$ then

$$M(x) \succeq \begin{bmatrix} Z & 0 \\ 0 & 0 \end{bmatrix} \iff (M/M_{22})(x) \succeq Z. \tag{20}$$

With these preliminaries established, we turn to the proof of Proposition 3. First suppose there is $Z \in \mathbb{S}_+^{n-1,(k)}$ such that $M(x) - \begin{bmatrix} Z & 0 \\ 0 & 0 \end{bmatrix} \succeq 0$. There are two cases to consider, depending on whether $M_{22}(x)$ is positive or zero.

Suppose we are in the case where $e_1(x) = nM_{22}(x) > 0$. Then $(M/M_{22})(x) \succeq Z$, so there is some $Z' \in \mathbb{S}_+^{n-1}$ such that

$$(M/M_{22})(x) = Z + Z' \in \mathbb{S}_+^{n-1,(k)} + \mathbb{S}_+^{n-1} = \mathbb{S}_+^{n-1,(k)}$$

where the last equality holds because $\mathbb{S}_+^{n-1,(k)} \supset \mathbb{S}_+^{n-1}$. It follows that $x \in \mathbb{R}_+^{n,(k)}$ because $e_1(x) > 0$ (by assumption) and by Lemma 4,

$$ie_i(x) = e_1(x)E_{i-1}((M/M_{22})(x)) \geq 0 \quad \text{for } i = 2, 3, \ldots, n-k.$$

Now consider the case where $e_1(x) = nM_{22}(x) = 0$. Since

$$\begin{bmatrix} M_{11}(x) - Z & M_{12}(x) \\ M_{12}(x)^T & M_{22}(x) \end{bmatrix} = \begin{bmatrix} M_{11}(x) - Z & V_n^T x/\sqrt{n} \\ x^T V_n/\sqrt{n} & 0 \end{bmatrix} \succeq 0$$

it follows that $V_n^T x = 0$. Since, $\hat{\mathbf{1}}_n^T x = 0$ we see that $Q_n^T x = 0$ so $x = 0 \in \mathbb{R}_+^{n,(k)}$.

Consider the reverse inclusion and suppose $x \in \mathbb{R}_+^{n,(k)}$. Again there are two cases depending on whether $e_1(x)$ is positive or zero. If $e_1(x) > 0$ take $Z = (M/M_{22})(x)$. Then, by (20), $M(x) \succeq \left[\begin{smallmatrix} Z & 0 \\ 0 & 0 \end{smallmatrix}\right]$. To see that $Z \in \mathbb{S}_+^{n-1,(k)}$ note that by Lemma 4,

$$E_i((M/M_{22})(x)) = (i+1)\frac{e_{i+1}(x)}{e_1(x)} \geq 0 \quad \text{for } i = 1, 2, \ldots, n-1-k.$$

If $x \in \mathbb{R}_+^{n,(k)}$ and $e_1(x) = 0$ then we use the assumption that $k \leq n-2$. Under this assumption $x \in \mathbb{R}_+^{n,(k)} \cap \{x : e_1(x) = 0\} = \{0\}$. In this case we can simply take $Z = 0 \in \mathbb{S}_+^{n-1,(k)}$ since $M(x) = 0 \succeq 0 = \left[\begin{smallmatrix} Z & 0 \\ 0 & 0 \end{smallmatrix}\right]$. $\qquad\square$

3.3   Dual relationships

We conclude this section by establishing Propositions 2D and 3D, the dual versions of Propositions 2 and 3. Both follow from general results about conic duality, such as the following rephrasing of [19, Corollary 16.3.2].

Lemma 5. *Suppose $K \subset \mathbb{R}^m$ is a closed convex cone and $A : \mathbb{R}^p \to \mathbb{R}^m$ and $B : \mathbb{R}^p \to \mathbb{R}^n$ are linear maps. Let*

$$C = \{B(x) : A(x) \in K\} \subset \mathbb{R}^n.$$

Furthermore, assume that there is $x_0 \in \mathbb{R}^p$ such that $A(x_0)$ is in the relative interior of K. Then

$$C^* = \{w \in \mathbb{R}^n : \exists y \in K^* \ \ s.t. \ \ B^*(w) = A^*(y)\}.$$

of Proposition 2D. Define $A : \mathbb{R}^n \to \mathbb{S}^{n-1}$ by $A(x) = V_n^T \operatorname{diag}(x) V_n$ and define B to be the identity on $\mathbb{R}^n$. Then by Proposition 2

$$\mathbb{R}_+^{n,(k)} = \{B(x) : A(x) \in \mathbb{S}_+^{n-1,(k-1)}\}.$$

Clearly B^* is the identity on $\mathbb{R}^n$ and $A^* : \mathbb{S}^{n-1} \to \mathbb{R}^n$ is given by $A^*(Y) = \operatorname{diag}(V_n Y V_n^T)$. Since $A(\mathbf{1}_n) = I_{n-1}$ is in the interior of $\mathbb{S}_+^{n-1,(k-1)}$, applying Lemma 5 we obtain

$$(\mathbb{R}_+^{n,(k)})^* = \{w \in \mathbb{R}^n : \exists Y \in (\mathbb{S}_+^{n-1,(k-1)})^* \ \ s.t. \ \ w = \operatorname{diag}(V_n Y V_n^T).\}$$

Eliminating w gives the statement in Proposition 2D. $\qquad\square$

of Proposition 3D. Define $A : \mathbb{R}^n \times \mathbb{S}^{n-1} \to \mathbb{S}^n \times \mathbb{S}^{n-1}$ by

$$A(x, Z) = (\operatorname{diag}(x) - V_n Z V_n^T, Z)$$

and $B : \mathbb{R}^n \times \mathbb{S}^{n-1} \to \mathbb{R}^n$ by $B(x, Z) = x$. Then by Proposition 3

$$\mathbb{R}_+^{n,(k)} = \{B(x, Z) : A(x, Z) \in \mathbb{S}_+^n \times \mathbb{S}_+^{n-1,(k)}\}.$$

A straightforward computation shows that $B^* : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{S}^{n-1}$ is given by $B^*(w) = (w, 0)$. Furthermore $A^* : \mathbb{S}^n \times \mathbb{S}^{n-1}$ is given by $A^*(Y, W) = (\operatorname{diag}(Y), W - V_n^T Y V_n)$. Since $A(2\mathbf{1}_n, I_{n-1})$ is in the interior of $\mathbb{S}_+^n \times \mathbb{S}_+^{n-1,(k)}$, applying Lemma 5 we obtain

$$(\mathbb{R}_+^{n,(k)})^* = \{w \in \mathbb{R}^n : \exists (Y, W) \in \mathbb{S}_+^n \times (\mathbb{S}_+^{n-1,(k)})^* \ \ s.t. \ \ w = \operatorname{diag}(W), \ \ V_n^T Y V_n = W\}.$$

Eliminating W and w gives the statement in Proposition 3D. $\qquad\square$

4 Exploiting symmetry: relating $\mathbb{S}_+^{n,(k)}$ and $\mathbb{R}_+^{n,(k)}$ and their dual cones

In the introduction we observed that $\mathbb{S}_+^{n,(k)}$ is invariant under the action of orthogonal matrices by conjugation on $\mathbb{S}^n$ and that its diagonal slice is $\mathbb{R}_+^{n,(k)}$. In this section we explain how to use these properties to construct the semidefinite representation of $\mathbb{S}_+^{n,(k)}$ in terms of $\mathbb{R}_+^{n,(k)}$ stated in Proposition 1. We then discuss how the duals of these two cones relate. The material in this section is well known so in some places we give appropriate references to the literature rather than providing proofs.

Let $O(n)$ denote the group of $n \times n$ orthogonal matrices. The *Schur-Horn cone* is

$$\mathrm{SH}_n = \left\{ (X, z) : \ z_1 \geq z_2 \geq \cdots \geq z_n, \ X \in \mathrm{conv}_{Q \in O(n)} \{ Q^T \mathrm{diag}(z) Q \} \right\}, \tag{21}$$

the set of pairs (X, z) such that z is in weakly decreasing order and X is in the convex hull of symmetric matrices with ordered spectrum z. We call this the Schur-Horn cone because all *symmetric Schur-Horn orbitopes* [20] appear as slices of SH_n of the form $\{ X : (X, z_0) \in \mathrm{SH}_n \}$ where z_0 is fixed and in weakly decreasing order.

Whenever a convex subset $C \subset \mathbb{S}^n$ is invariant under orthogonal conjugation, i.e. C is a spectral set, we can express C in terms of the Schur-Horn cone and the (hopefully simpler) diagonal slice of C as follows.

Lemma 6. *If $C \subset \mathbb{S}^n$ is convex and invariant under orthogonal conjugation then*

$$C = \{ X \in \mathbb{S}^n : \exists z \in \mathbb{R}^n \ s.t. \ (X, z) \in \mathrm{SH}_n, \ \mathrm{diag}(z) \in C \}.$$

Proof. Suppose $X \in C$. Take $z = \lambda(X)$, the ordered vector of eigenvalues of X. Then there is some $Q \in O(n)$ such that $X = Q^T \mathrm{diag}(\lambda(X)) Q$ so $(X, \lambda(X)) \in \mathrm{SH}_n$. By the orthogonal invariance of C, $X \in C$ implies that $QXQ^T = \mathrm{diag}(\lambda(X)) \in C$.

For the reverse inclusion, suppose there is $z \in \mathbb{R}^n$ such that $(X, z) \in \mathrm{SH}_n$ and $\mathrm{diag}(z) \in C$. Then by the orthogonal invariance of C, $Q^T \mathrm{diag}(z) Q \in C$ for all $Q \in O(n)$. Since C is convex, $\mathrm{conv}_{Q \in O(n)} \{ Q^T \mathrm{diag}(z) Q \} \subseteq C$. Hence $(X, z) \in \mathrm{SH}_n$ implies that

$$X \in \mathrm{conv}_{Q \in O(n)} \{ Q^T \mathrm{diag}(z) Q \} \subseteq C.$$

$\square$

The first statement in Proposition 1 follows from Lemma 6 by recalling that $\mathbb{S}_+^{n,(k)}$ is orthogonally invariant and $\mathbb{R}_+^{n,(k)} = \{ z \in \mathbb{R}^n : \mathrm{diag}(z) \in \mathbb{S}_+^{n,(k)} \}$.

Proving the remainder of Proposition 1 then reduces to establishing the correctness of the stated semidefinite representation of SH_n. This can be deduced from the following two well-known results.

Lemma 7. *If $\lambda(X)$ is ordered so that $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$ then $(X, z) \in \mathrm{SH}_n$ if and only if $z_1 \geq z_2 \geq \cdots \geq z_n$,*

$$\mathrm{tr}(X) = \sum_{i=1}^n \lambda_i(X) = \sum_{i=1}^n z_i, \quad and \quad \sum_{i=1}^\ell \lambda_i(X) \leq \sum_{i=1}^\ell z_i \quad for \ \ell = 1, 2, \ldots, n-1.$$

In other words $(X, z) \in \mathrm{SH}_n$ if and only if z is weakly decreasing and $\lambda(X)$ is *majorized* by z. This is discussed, for example, in [20, Corollary 3.2]. To turn this characterization into a semidefinite representation, it suffices to have semidefinite representations of the epigraphs of the convex functions $s_\ell(X) := \sum_{i=1}^\ell \lambda_i(X)$. These are given by Nesterov and Nemirovski in [15, Section 6.4.3, Example 7].

Lemma 8. *If $2 \leq \ell \leq n - 1$, the epigraph of the convex function $s_\ell(X) = \sum_{i=1}^{\ell} \lambda_i(X)$ has a semidefinite representation of size $O(n)$ given by*

$$\{(X,t) : s_\ell(X) \leq t\} = \{(X,t) : \exists s \in \mathbb{R},\ Z \in \mathbb{S}^n \quad s.t. \quad Z \succeq 0,\ X \preceq Z + sI,\ \text{tr}(Z) + s\ell \leq t\}.$$

The epigraph of $s_1(X)$ has a simpler semidefinite representation as

$$\{(X,t) : s_1(X) \leq t\} = \{(X,t) : X \preceq tI\}.$$

We now turn to the relationship between $(\mathbb{S}_+^{n,(k)})^*$ and $(\mathbb{R}_+^{n,(k)})^*$. Note that $(\mathbb{S}_+^{n,(k)})^*$ is invariant under orthogonal conjugation. So the claim (Proposition 1D) that

$$(\mathbb{S}_+^{n,(k)})^* = \{Y \in \mathbb{S}^n : \exists w \in \mathbb{R}^n \ \text{s.t.} \ w \in (\mathbb{R}_+^{n,(k)})^*,\ (Y, w) \in \text{SH}_n\}$$

would follow from Lemma 6 once we know that the diagonal slice of $(\mathbb{S}_+^{n,(k)})^*$ is $(\mathbb{R}_+^{n,(k)})^*$. This is a special case of the following result for which we give a direct proof.

Lemma 9. *Suppose $C \subset \mathbb{S}^n$ is a convex cone that is invariant under orthogonal conjugation. Then*

$$\{y \in \mathbb{R}^n : \text{diag}(y) \in C^*\} = \{z \in \mathbb{R}^n : \text{diag}(z) \in C\}^*. \tag{22}$$

Note that if $C = \mathbb{S}_+^{n,(k)}$ then the left hand side of (22) is the diagonal slice of $(\mathbb{S}_+^{n,(k)})^*$ and the right hand side is $(\mathbb{R}_+^{n,(k)})^*$.

Proof. We use a description of the orthogonal projector onto the subspace of diagonal matrices as an average of orthogonal conjugations (see, e.g., [12] where the idea is attributed to Olkin). For every subset $I \subset \{1, 2, \ldots, n\}$ let Δ_I denote the diagonal matrix with $[\Delta_I]_{ii} = 1$ if $i \in I$ and $[\Delta_I]_{ii} = -1$ otherwise. The Δ_I are all orthogonal and act on symmetric matrices by $X \mapsto \Delta_I X \Delta_I^T$. A symmetric matrix is fixed by the action of all the Δ_I if and only if it is diagonal. Hence $\text{diag}(\text{diag}(X))$, the orthogonal projection of a symmetric matrix X onto the subspace of diagonal matrices (the fixed-point subspace), is given by averaging over the action of the Δ_I, i.e.

$$\text{diag}(\text{diag}(X)) = \frac{1}{2^n} \sum_I \Delta_I X \Delta_I^T \tag{23}$$

where the sum is over all 2^n subsets of $\{1, 2, \ldots, n\}$.

We now prove that $\{\text{diag}(X) : X \in C\} = \{x \in \mathbb{R}^n : \text{diag}(x) \in C\}$. Observe that the diagonal slice of C is certainly contained in the diagonal projection of C giving one inclusion. For the other, suppose $X \in C$ is arbitrary. Since C is orthogonally invariant, each $\Delta_I X \Delta_I^T$ is an element of C. Since C is convex, it follows from (23) that $\text{diag}(\text{diag}(X)) \in C$ and is diagonal as we require.

To prove (22), we apply Lemma 5 in Section 3.3 to obtain an expression for $\{x \in \mathbb{R}^n : \text{diag}(x) \in C\}^*$. For Lemma 5 to apply, we must exhibit $x_0 \in \mathbb{R}^n$ such that $\text{diag}(x_0)$ is in the relative interior of C, denoted $\text{relint}(C)$. Let $X_0 \in \text{relint}(C)$ be arbitrary. Since C is invariant under orthogonal conjugation, the same holds for $\text{relint}(C)$. It follows that each $\Delta_I X_0 \Delta_I^T \in \text{relint}(C)$ and by (23) (and the convexity of $\text{relint}(C)$) it follows that $\text{diag}(\text{diag}(X_0)) \in \text{relint}(C)$. As such it suffices to take $x_0 = \text{diag}(X_0)$. □

5 Concluding remarks

We conclude with some comments about (the possibility of) simplifying our representations and some open questions.

5.1 Simplifications

If we can simplify a representation of $\mathbb{R}_+^{n,(k)}$ or $\mathbb{S}_+^{n,(k)}$ for some $k = i$, that allows us to simplify the derivative-based representations for $k \geq i$ and the polar derivative-based representations for $k \leq i$. For example $\mathbb{R}_+^{n,(n-2)}$ can be succinctly expressed in terms of the second-order cone $Q_+^{n+1} = \{x \in \mathbb{R}^{n+1} : (\sum_{i=1}^n x_i^2)^{1/2} \leq x_{n+1}\}$ as

$$\mathbb{R}_+^{n,(n-2)} = \{x \in \mathbb{R}^n : (x, e_1(x)) \in Q_+^{n+1}\}.$$

Then we can represent $\mathbb{S}_+^{n,(n-2)}$ in terms of the second-order cone as

$$\mathbb{S}_+^{n,(n-2)} = \{Z \in \mathbb{S}^n : (Z, \operatorname{tr}(Z)) \in Q_+^{n^2+1}\}$$

because $\operatorname{tr}(Z) = \sum_{i=1}^n \lambda_i(Z)$ and $\sum_{i,j=1}^n Z_{ij}^2 = \sum_{i=1}^n \lambda_i(Z)^2$. This should be used as a base case instead of $\mathbb{S}_+^{n,(n-1)}$ in the polar derivative-based representations.

As an example of this, Proposition 3 can be used to give a concise representation of $\mathbb{R}_+^{n,(n-3)}$ in terms of the second-order cone as

$$x \in \mathbb{R}_+^{n,(n-3)} \quad \Longleftrightarrow \quad \exists Z \in \mathbb{S}^{n-1} \text{ such that}$$

$$\operatorname{diag}(x) \succeq V_n Z V_n^T \text{ and } (Z, \operatorname{tr}(Z)) \in Q_+^{(n-1)^2+1}.$$

5.2 Lower bounds on the size of representations

The explicit constructions given in this paper establish upper bounds on the minimum size of semidefinite representations of $\mathbb{S}_+^{n,(k)}$ and $\mathbb{R}_+^{n,(k)}$. To assess how good our representations are, it is interesting to establish corresponding *lower* bounds on the size of semidefinite representations of $\mathbb{R}_+^{n,(k)}$ and $\mathbb{S}_+^{n,(k)}$. Since $\mathbb{R}_+^{n,(k)}$ is a slice of $\mathbb{S}_+^{n,(k)}$, any lower bound on the size of a semidefinite representation of $\mathbb{R}_+^{n,(k)}$ also provides a lower bound on the size of a semidefinite representation of $\mathbb{S}_+^{n,(k)}$. Hence we focus our discussion on $\mathbb{R}_+^{n,(k)}$.

In the case of $\mathbb{R}_+^{n,(n-1)}$, a halfspace, the obvious semidefinite representation of size one is clearly of minimum size. Less trivial is the case of $\mathbb{R}_+^{n,(0)}$, the non-negative orthant. It has been shown by Gouveia et al. [5, Section 5] that $\mathbb{R}_+^n$ does not admit a semidefinite representation of size smaller than n. Hence the obvious representation of $\mathbb{R}_+^n$ as the restriction of $\mathbb{S}_+^n$ to the diagonal is of minimum size.

For each k, the slice of $\mathbb{R}_+^{n,(k)}$ obtained by setting the last k variables to zero is $\mathbb{R}_+^{n-k}$. Hence any semidefinite representation of $\mathbb{R}_+^{n,(k)}$ has size at least $n - k$, the minimum size of a semidefinite representation of $\mathbb{R}_+^{n-k}$. This argument establishes that Sanyal's spectrahedral representation of $\mathbb{R}_+^{n,(1)}$ of size $n - 1$ is actually a minimum size semidefinite representation of $\mathbb{R}_+^{n,(1)}$. We are not aware of any other lower bounds on the size of semidefinite representations of the cones $\mathbb{R}_+^{n,(k)}$ for $2 \leq k \leq n - 2$.

The semidefinite representations of $\mathbb{R}_+^{n,(k)}$ given in this paper are *equivariant* in that they appropriately preserve the symmetries of $\mathbb{R}_+^{n,(k)}$. (For a precise definition see [5, Definition 4].) It is known that symmetry matters when representing convex sets as projections of other convex sets [9]. For example if p is a power of a prime, equivariant representations of regular p-gons in $\mathbb{R}^2$ are necessarily much larger than their minimum-sized non-equivariant counterparts [5, Proposition 3]. Given that the cones $\mathbb{R}_+^{n,(k)}$ are highly symmetric, it would also be interesting to establish lower bounds on the size of equivariant semidefinite representations of the derivative relaxations of the non-negative orthant.

References

[1] G. Blekherman, P. A. Parrilo, and R. R. Thomas, editors. *Semidefinite Optimization and Convex Algebraic Geometry.* MOS-SIAM Series on Optimization. SIAM, Philadelphia, 2013.

[2] P. Bränden. Hyperbolicity cones of elementary symmetric polynomials are spectrahedral. *Optim. Lett.*, 8(5):1773–1782, 2014.

[3] Y.-B. Choe, J. G. Oxley, A. D. Sokal, and D. G. Wagner. Homogeneous multivariate polynomials with the half-plane property. *Adv. Appl. Math.*, 32(1–2):88–187, 2004.

[4] J. Gouveia and T. Netzer. Positive polynomials and projections of spectrahedra. *SIAM J. Optim.*, 21(3):960, 2011.

[5] João Gouveia, Pablo A Parrilo, and Rekha R Thomas. Lifts of convex sets and cone factorizations. *Math. Oper. Res.*, 38(2):248–264, 2013.

[6] L. Gårding. An inequality for hyperbolic polynomials. *J. Math. Mech*, 8(6):957–965, 1959.

[7] O. Güler. Hyperbolic polynomials and interior point methods for convex programming. *Math. Oper. Res.*, 22(2):350–377, 1997.

[8] J. W. Helton and V. Vinnikov. Linear matrix inequality representation of sets. *Comm. Pure Appl. Math*, 60(5):654–674, 2007.

[9] V. Kaibel, K. Pashkovich, and D. Theis. Symmetry matters for the sizes of extended formulations. In *Integer Programming and Combinatorial Optimization*, volume 6080 of *Lecture Notes in Comput. Sci.*, pages 135–148. Springer Berlin, 2010.

[10] A. S. Lewis, P. A. Parrilo, and M. V. Ramana. The Lax conjecture is true. *Proc. Amer. Math. Soc.*, 133(9):2495–2500, 2005.

[11] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[12] H. F. Miranda and R. C. Thompson. Group majorization, the convex hulls of sets of matrices, and the diagonal element-singular value inequalities. *Linear Alg. Appl.*, 199:131–141, 1994.

[13] A. Nemirovski. Advances in convex optimization: conic programming. In *Proceedings of the International Congress of Mathematicians: Madrid, August 22–30, 2006: invited lectures*, pages 413–444, 2006.

[14] A. Nemirovski and A. Ben-Tal. *Lectures on Modern Convex Optimization: Analysis, Algorithms and Engineering Applications.* MOS-SIAM Series on Optimization. SIAM, 2001.

[15] Yu. Nesterov and A. Nemirovskii. *Interior Point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM Studies in Applied Mathematics.* SIAM, Philadelphia, 1993.

[16] J. Nie, P. A. Parrilo, and B. Sturmfels. Semidefinite representation of the k-ellipse. In *Algorithms in Algebraic Geometry*, volume 146 of *The IMA Volumes in Mathematics and its Applications*, pages 117–132. Springer New York, 2008.

[17] M. Ramana and A. J. Goldman. Some geometric results in semidefinite programming. *J. Global Optim.*, 7(1):33–50, 1995.

[18] J. Renegar. Hyperbolic programs, and their derivative relaxations. *Found. Comput. Math.*, 6:59–79, 2006.

[19] R. T. Rockafellar. *Convex analysis*, volume 28. Princeton university press, 1997.

[20] R. Sanyal, F. Sottile, and B. Sturmfels. Orbitopes. *Mathematika*, 57(02):275–314, 2011.

[21] Raman Sanyal. On the derivative cones of polyhedral cones. *Adv. Geom.*, 13(2):315–321, 2013.

[22] K. C. Toh, M. J. Todd, and R. H. Tütüncü. SDPT3—a MATLAB software package for semidefinite programming, version 1.3. *Optim. Methods Softw.*, 11(1-4):545–581, 1999.

[23] Y. Zinchenko. On hyperbolicity cones associated with elementary symmetric polynomials. *Optim. Lett.*, 2(3):389–402, 2008.