A lower bound on the positive semidefinite rank of convex bodies

Hamza Fawzi[*] Mohab Safey El Din[†]

January 8, 2018

Abstract

The positive semidefinite rank of a convex body C is the size of its smallest positive semidefinite formulation. We show that the positive semidefinite rank of any convex body C is at least $\sqrt{\log d}$ where d is the smallest degree of a polynomial that vanishes on the boundary of the polar of C. This improves on the existing bound which relies on results from quantifier elimination. Our proof relies on the Bézout bound applied to the Karush-Kuhn-Tucker conditions of optimality. We discuss the connection with the algebraic degree of semidefinite programming and show that the bound is tight (up to constant factor) for random spectrahedra of suitable dimension.

1 Introduction

Semidefinite programming is the problem of optimizing a linear function over a convex set described by a linear matrix inequality:

$$\max \quad c^T x \quad \text{s.t.} \quad x \in S$$

where $S \subseteq \mathbb{R}^n$ has the form:

$$S = \{x \in \mathbb{R}^n : A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0\}. \tag{1}$$

Here $A_0, \dots, A_n$ are real symmetric matrices of size $m \times m$ and the notation $M \succeq 0$ indicates that M is positive semidefinite. A set of the form (1) is called a *spectrahedron*.

Given a convex set $C \subseteq \mathbb{R}^k$, we say that C has a *semidefinite lift* of size m if it can be expressed as

$$C = \pi(S)$$

where S is a spectrahedron (1) defined using matrices of size $m \times m$ and π is any linear map. If C can be expressed in this way, then any linear optimization problem over C can be expressed as a semidefinite program of size m. The size of the smallest semidefinite lift of C is called the *positive semidefinite rank* of C [GPT13, FGP$^+$15].

The purpose of this paper is to give a general lower bound on the positive semidefinite rank of convex bodies. Here, by a convex body we mean a closed convex set such that the origin lies in the interior of C. For the statement of our main theorem, we need the notion of *polar* of a convex body C, defined as follows:

$$C^o = \left\{ c \in \mathbb{R}^k : \langle c, x \rangle \leq 1 \; \forall x \in C \right\}.$$

[*] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, UK.

[†] Sorbonne Universités, UPMC Univ Paris 06, CNRS, INRIA Paris Center, LIP6, Equipe PolSys, F-75005, Paris, France.

The polar of a convex body is another full-dimensional closed convex set that is bounded and contains the origin [Roc97, Theorem 14.6]. Throughout the paper, we use log for the logarithm base 2. We can now state the first main result of this article.

Theorem 1. *Let C be a convex body and let C^o be its polar. Let d be the smallest degree of a polynomial with real coefficients that vanishes on the boundary of C^o. Then* $\operatorname{rank}_{\text{psd}}(C) \geq \sqrt{\log d}$.

We also show that this bound is tight in general (up to multiplicative factors):

Theorem 2. *There exist convex bodies C where the degree d of the algebraic boundary of C^o can be made arbitrary large and where* $\operatorname{rank}_{\text{psd}}(C) \leq \sqrt{20 \log d}$.

When C is a polytope, the degree d of the algebraic boundary of C^o is nothing but the number of vertices of C. Theorem 1 can thus be compared to the well-known lower bound of Goemans [Goe15] on the size of linear programming lifts. The *linear programming extension complexity* of a polytope P is the smallest f such that P can be written as the linear projection of a polytope with f facets.

Theorem 3 (Goemans [Goe15, Theorem 1])**.** *Assume P is a polytope with d vertices. Then the linear programming extension complexity of P is at least* $\log d$.

Proof. The proof is elementary so we include it for completeness. Assume $P = \pi(Q)$ where Q is a polytope with f facets. The pre-image by π of any vertex of P is a face of Q. Since Q has at most 2^f faces it follows that $f \geq \log d$. □

For functions $f, g : \mathbb{N} \to \mathbb{R}$ we say that $f(n) \in \Omega(g(n))$ if there exists a constant $K > 0$ such that $f(n) \geq K \cdot g(n)$ for all large enough n.

The only previous lower bound on the positive semidefinite rank that applies to arbitrary convex bodies that we are aware of is the bound proved in [GPT13, Proposition 6] which says that[1] $\operatorname{rank}_{\text{psd}}(C) \geq \Omega\left(\sqrt{\frac{\log d}{n \log\log(d/n)}}\right)$ where n is the dimension of C. This bound was obtained using results from quantifier elimination theory and one drawback is that it involves constants that are unknown or difficult to estimate. Our lower bound of Theorem 1 improves on this existing bound and also has the advantage of being explicit.

Main ideas. The main idea behind the proof of Theorem 1 is simple. Given a convex body C, we exhibit a system of polynomial equations that vanishes on the boundary of C^o. This system of polynomial equations is nothing but the Karush-Kuhn-Tucker (KKT) system, after discarding the inequality constraints to get an algebraic variety. Applying the Bézout theorem on the KKT system gives us an upper bound on the degree of this variety and yields the stated lower bound. To prove Theorem 2 about tightness of the bound we appeal to existing works [NRS10] where the degree of the KKT system was explicitly computed, under certain genericity assumptions. The convex bodies of Theorem 2 are in fact random spectrahedra (i.e., spectrahedra defined using random matrices $A_0, \ldots, A_n$) of appropriate dimension, where the formulas for the algebraic degree of semidefinite programming [vBR09] allow us to lower bound the degree of the algebraic boundary of their polars. We would like to point out that many of the ideas involved in the proofs of Theorems 1 and 2 appear in some form or another in [RS12, NRS10, SS15]. For example a study of the algebraic boundary of polars of spectrahedra appears in [RS12, Section 5.5]. However it seems that the connection

[1] In the bound shown in [GPT13, Proposition 6], d is the degree of the algebraic boundary of C. However since $\operatorname{rank}_{\text{psd}}(C) = \operatorname{rank}_{\text{psd}}(C^o)$ it can also be taken to be that of C^o in the statement of the lower bound.

with the positive semidefinite rank was not made explicit before. The focus in these previous works seemed to be on getting *exact* values for the degrees, at the price of genericity assumptions. In the present work our aim was on getting bounds (tight up to constant factors) but valid without any genericity assumption.

Notations. The (topological) boundary of a set $C \subseteq \mathbb{R}^n$ is denoted ∂C and defined as $\partial C = \mathbf{cl}(C) \setminus \mathbf{int}(C)$ where **cl** and **int** denote closure and interior respectively. The algebraic boundary of $C \subseteq \mathbb{R}^n$ denoted $\partial_a C$ is the smallest affine algebraic variety in $\mathbb{C}^n$ that contains ∂C. We denote by $\mathbf{S}^m$ the space of $m \times m$ real symmetric matrices. This is a real vector space of dimension

$$t_m := \binom{m+1}{2}.$$

We also denote by $\mathbf{S}^m(\mathbb{C})$ the space of $m \times m$ symmetric matrices with complex entries.

2 Proof of Theorem 1

In this section we prove Theorem 1. To do so we will exhibit polynomial equations that vanish on the boundary of polars of spectrahedra and their shadows. These equations are nothing but the KKT conditions of optimality. Applying the Bézout bound will yield Theorem 1.

KKT equations. Let $A_0, \ldots, A_n \in \mathbf{S}^m$ and define

$$\mathcal{A}(x) := x_1 A_1 + \cdots + x_n A_n.$$

Consider the linear optimization problem

$$\max \quad c^T x \quad \text{s.t.} \quad A_0 + \mathcal{A}(x) \succeq 0 \tag{2}$$

and assume that the feasible set

$$S = \{x \in \mathbb{R}^n : A_0 + \mathcal{A}(x) \succeq 0\}$$

contains 0 in its interior. In this case we know that any optimal point x of (2) must satisfy the following KKT conditions:

$$\exists X, Z \in \mathbf{S}^m \quad : \quad X = A_0 + \mathcal{A}(x), \quad \mathcal{A}^*(Z) + c = 0, \quad XZ = 0, \quad X \succeq 0, \quad Z \succeq 0 \tag{3}$$

where the variable Z plays the role of *dual multiplier* and $\mathcal{A}^*(Z) = (\mathsf{Trace}(A_1\, Z), \ldots, \mathsf{Trace}(A_n\, Z))$. Conditions (3) consist of equality conditions as well as inequality conditions. If we disregard the inequality conditions we get a system of polynomial equations in $(x, X, Z) \in \mathbb{R}^n \times \mathbf{S}^m \times \mathbf{S}^m$ which we denote by KKT(c):[2]

$$\text{KKT}(c) \quad : \quad X = A_0 + \mathcal{A}(x), \quad \mathcal{A}^*(Z) + c = 0, \quad XZ = 0. \tag{4}$$

This system has $n + 2t_m$ unknowns and consists of $n + t_m + m^2$ equations. A crucial fact about this system is that it has a finite number of solutions, assuming the parameters $A_0, A_1, \ldots, A_n$ and

[2] We note here that there are multiple ways of writing the SDP complementarity conditions in general, and these can lead to differences in the context of algorithms for SDP, see e.g., the discussion in [BTN01, Section 6.5.4]. For our purposes, the main property that we will need of the system KKT(c) is that it has a finite number of solutions generically (Lemma 1).

c are generic (we come back to the genericity assumption after the statement of the result; some form of genericity is needed for the statement to be true). It is the number of solutions to the KKT system that will give an upper bound on the degree of the algebraic boundary of the polar as we will see later.

Lemma 1 (Finiteness of KKT solutions)**.** *For generic $A_0, A_1, \ldots, A_n$ and c, the KKT system of polynomial equations (4) has a finite number of complex solutions $(x, X, Z) \in \mathbb{C}^n \times \mathbf{S}^m(\mathbb{C}) \times \mathbf{S}^m(\mathbb{C})$. Furthermore the number of such solutions is at most 2^{m^2}.*

Proof. That the KKT system has a finite number of solutions generically was proved in [NRS10, Theorem 7]. We include a sketch of proof for completeness which is simply a dimension count argument. There are three equations in (4):

- The equation $X = A_0 + \mathcal{A}(x)$ is linear and defines an affine subspace of codimension t_m (we assume that $\mathcal{A}$ is injective).
- The equation $\mathcal{A}^*(Z) + c = 0$ is also linear and defines an affine subspace of codimension n.
- Finally the equations $XZ = 0$ can be shown to define a variety of codimension t_m (see e.g., [NRS10, Proof of Theorem 7]).

If $A_0, A_1, \ldots, A_n$ and c are generic, a Bertini-Sard type theorem tells us that the intersection of these three varieties will have codimension equal to the sum of the codimensions, i.e., $t_m + n + t_m = 2t_m + n$ which is the dimension of the ambient space. In other words the variety defined by (4) is zero-dimensional, i.e., there are a finite number of solutions.

Bézout bound tells us that the number of solutions is at most the product of the degrees of the polynomial equations that form the system (4), which in this case is 2^{m^2}. □

Remark 1 (Genericity assumption of Lemma 1)**.** *An assumption of genericity is necessary in general to guarantee that the system (4) has a finite number of solutions. This is to rule out situations where the optimization problem (2) has an infinite number of solutions (a positive dimensional face of S) or when there are an infinite number of dual multipliers. In Lemma 1 we assumed all the parameters $A_0, \ldots, A_n, c$ generic to be able to apply a standard Bertini-Sard type theorem. We think however it may be possible to remove some of the genericity assumptions (e.g., just to assume genericity on A_0 and c) but we did not pursue this further here as the current statement of the lemma will be sufficient for our purposes.*

The next lemma shows that the number of solutions to the KKT system is intimately tied to the degree of the algebraic boundary of the polar S^o.

Lemma 2. *Consider a spectrahedron $S = \{x \in \mathbb{R}^n : A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0\}$ where $A_0, \ldots, A_n \in \mathbf{S}^m$ and assume that $0 \in \textbf{\textit{int}}\, S$. Let S^o be its polar defined by*

$$S^o = \{c \in \mathbb{R}^n : \langle c, x \rangle \leq 1 \ \forall x \in S\}.$$

Then there is a polynomial of degree at most 2^{m^2} with real coefficients that vanishes on the boundary of S^o.

Proof. The points on the boundary of S^o are exactly those c such that $\max_{x \in S} c^T x = 1$. Consider the system of polynomial equations obtained by adding the equation $c^T x = 1$ to the KKT system:

$$\text{KKT} \ : \ \begin{cases} X = A_0 + \mathcal{A}(x), \quad \mathcal{A}^*(Z) + c = 0, \quad XZ = 0 \\ c^T x = 1. \end{cases} \tag{5}$$

We think of (5) as a system of equations on the variables (x, X, Z, c). If we eliminate the variables x, X, Z we get an algebraic variety $\mathcal{V} \subset \mathbb{C}^n$ in the variables c:

$$\mathcal{V} = \mathrm{elim}_c(\mathsf{Sols}(\mathrm{KKT})). \tag{6}$$

By construction this variety contains the boundary of S^o, i.e., $\partial S^o \subseteq \mathcal{V} \cap \mathbb{R}^n$. To bound the degree of $\partial_a S^o$ it thus suffices to count the number of intersections of $\mathcal{V}$ with a generic line, since $\partial_a S^o \subseteq \mathcal{V}$ and $\partial_a S^o$ is a hypersurface [Sin15, Corollary 2.8]. We will do this first in the case where $A_0, \dots, A_n$ are generic. Let $c_0 \in \mathbb{C}^n$ generic and consider the line $\{\lambda c_0 : \lambda \in \mathbb{C}\}$. Since $\mathcal{V}$ was defined by eliminating variables x, X, Z from (5), we know that $\lambda c_0 \in \mathcal{V}$ if and only if there exist (x, X, Z) in the solution set of $\mathrm{KKT}(\lambda c_0)$ and $\lambda {c_0}^T x = 1$. By looking at the equations defining $\mathrm{KKT}(\lambda c_0)$ this implies that $(x, X, ({c_0}^T x)Z)$ is in the solution set of $\mathrm{KKT}(c_0)$. Thus the number of intersection points is at most the cardinality of the solution set of $\mathrm{KKT}(c_0)$, i.e., 2^{m^2}. We have thus shown that $\partial_a S^o$ is a hypersurface of degree at most 2^{m^2}.

It thus remains to treat the case where $A_0, A_1, \dots, A_n$ in the definition of S are not generic. This can be done by using a simple perturbation argument. Let N be the total number of the entries in $n+1$ symmetric matrices. Hence, the sequence of matrices $A_0, \dots, A_n$ represents a point $\mathbf{A}$ in $\mathbb{R}^N$. For any $k \in \mathbb{N} \setminus \{0\}$, there exists a point $\mathbf{A}_k$ in $\mathbb{R}^N$ in the ball centered at $\mathbf{A}$ of radius $1/k$ which is generic and represents a sequence of symmetric matrices $A_{0,k}, \dots, A_{n,k}$. Since, by assumption $0 \in \mathbf{int}\, S$, A_0 is positive definite, one can assume w.l.o.g. that $A_{0,k}$ is positive definite. Hence the spectrahedra S_k defined by $A_{0,k}, \dots, A_{n,k}$ are generic, non-empty and such that $0 \in \mathbf{int}\, S_k$. Hence, one can apply to them the above paragraph.

Now, let (p_k) be a sequence of polynomials of degree at most 2^{m^2} that vanish on the boundary of (S_k). We can rescale each p_k to be unit-normed and we can thus assume that (p_k) has a convergent subsequence that converges to some polynomial p. Clearly the degree of p is at most 2^{m^2}. Finally it is easy to verify that p vanishes on the boundary of S^o. □

We are now in position to prove Theorem 1 on the lower bound for the positive semidefinite rank. The main idea is that if $C = \pi(S)$ where S is a spectrahedron, then by duality C^o is the intersection of S^o with an affine subspace and thus the algebraic boundary of C^o has degree at most that of S^o.

Proof of Theorem 1. Assume C is a convex body that can be written as $C = \pi(S)$ where S is a spectrahedron defined using an $m \times m$ linear matrix inequality and π a linear map. We can assume that S has nonempty interior, and furthermore that $0 \in \mathbf{int}(S)$ since $0 \in \mathbf{int}(C)$. We are going to exhibit a polynomial of degree at most 2^{m^2} that vanishes on the boundary of C^o. Let p be a polynomial of degree at most 2^{m^2} that vanishes on the boundary of S^o. Then we claim that the polynomial $q = p \circ \pi^*$ (where π^* is the adjoint of π), which has degree at most 2^{m^2} vanishes on the boundary of C^o. Indeed if y is on the boundary of C^o this means that $\max_{x \in C} \langle y, x \rangle = 1$ which means that $\max_{x \in S} \langle \pi^*(y), x \rangle = 1$ and so $\pi^*(y)$ is on the boundary of S^o, hence $q(y) = p(\pi^*(y)) = 0$.

If we let d be the degree of the algebraic boundary of C^o and $m = \mathrm{rank}_{\mathrm{psd}}(C)$ we have thus shown that $d \leq 2^{m^2}$ which implies $\mathrm{rank}_{\mathrm{psd}}(C) = m \geq \sqrt{\log d}$. □

Application: number of vertices of spectrahedral shadows. In this subsection we discuss an application of Theorem 1 to bound the number of *vertices* of spectrahedral shadows. If $C \subseteq \mathbb{R}^n$ is a convex body and $x \in C$, the *normal cone* of C at x is defined as

$$N_C(x) := \{c \in \mathbb{R}^n : \langle c, z \rangle \leq \langle c, x \rangle \; \forall z \in C\}.$$

A point $x \in C$ is called a *vertex* if $N_C(x)$ is *full-dimensional*. Observe that any vertex of C must be an extreme point, but not all extreme points are vertices, see Figure 1. Vertices play the role of singularities on the boundaries of convex sets; in fact they are also sometimes called *0-singular points*. It is known, see e.g., [Sch13, Theorem 2.2.5] that any convex set has at most a countable number of vertices. Vertices of spectrahedra arising from combinatorial optimization problems have been studied in [LP95, dCST15]. The next theorem gives an upper bound on the number of vertices of any spectrahedral shadow. To the best of our knowledge this is the first such bound.

Theorem 4. *If C is a convex body having a semidefinite representation of size m, then C has at most 2^{m^2} vertices.*

Proof. Any vertex of C will contribute a linear factor in the algebraic boundary of C^o: indeed if x is a vertex of C then the algebraic boundary of C must contain the hyperplane $\{c \in \mathbb{R}^n : c^T x = 1\}$ (see e.g., Figure 1(right)). Thus the degree of $\partial_a C^o$ is greater than or equal the number of vertices of C. The result follows since the degree of $\partial_a C^o$ is at most 2^{m^2}. □

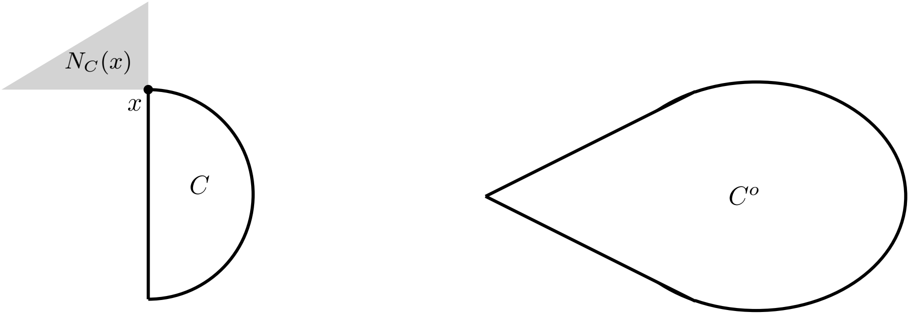

Figure 1: Left: A vertex of a convex set C. Right: The polar of C. We see that each vertex contributes a hyperplane in the algebraic boundary $\partial_a C^o$.

3 Tightness of lower bound, and algebraic degree of semidefinite programming

In this section we prove Theorem 2. We will show that the lower bound of Theorem 1 is tight up to a constant factor on certain random spectrahedra of appropriate dimension n, namely $n \approx t_{m/2}$.

Let S be a spectrahedron defined using matrices $A := (A_0, \ldots, A_n)$. In the previous section we saw that if we project the following KKT equations

$$\text{KKT} \; : \; \begin{cases} X = A_0 + \mathcal{A}(x), \quad \mathcal{A}^*(Z) + c = 0, \quad XZ = 0 \\ c^T x = 1 \end{cases} \tag{7}$$

on $c \in \mathbb{C}^n$ we get an algebraic variety

$$\mathcal{V}(A) = \text{elim}_c(\mathsf{Sols}(\text{KKT}))$$

that vanishes on the boundary of S^o. This variety could coincide exactly with $\partial_a S^o$ but it can also contain spurious components that do not intersect ∂S^o and thus are not in its Zariski closure (see Section 4 later for an example).

In order to prove our result we need to understand the irreducible components of the variety $\mathcal{V}(A)$. If we can show that there is an *irreducible component* $\mathcal{W}$ of $\mathcal{V}(A)$ whose intersection with the boundary of S^o has dimension the one of $\mathcal{W}$ then we know that the degree of the algebraic boundary of S^o is at least $\deg \mathcal{W}$. When A is generic, the irreducible components of $\mathcal{V}(A)$ have been studied in [NRS10] where it was shown that they are obtained by imposing rank conditions on the matrices X and Z in the KKT equations, namely by considering the following system for a fixed r:

$$\mathrm{KKT}_r \ : \ \begin{cases} X = \mathcal{A}(x) + A_0, \ \ \mathcal{A}^*(Z) + c = 0, \ \ XZ = 0 \\ c^T x = 1, \\ \operatorname{rank}(X) \leq r, \ \ \operatorname{rank}(Z) \leq m - r. \end{cases} \tag{8}$$

We think of (8) as a system of equations in (x, X, Z, c). If we eliminate the variables (x, X, Z) from the above equations we get an algebraic variety in $\mathbb{C}^n$ that is contained in $\mathcal{V}(A)$. We call this variety $\mathcal{V}_r(A)$:

$$\mathcal{V}_r(A) = \operatorname{elim}_c(\mathsf{Sols}(\mathrm{KKT}_r)) \subseteq \mathcal{V}(A). \tag{9}$$

For generic A, it was shown in [NRS10, Theorem 13] that $\mathcal{V}_r(A)$ is a hypersurface provided r satisfies the *Pataki bounds*:

$$n \geq t_{m-r} \qquad \text{and} \qquad t_r \leq t_m - n. \tag{10}$$

Using Bertini theorem one can show that this variety is also irreducible over $\mathbb{C}$ provided $n > t_{m-r}$.

Lemma 3. *For generic $A_0, \ldots, A_n$ the variety $\mathcal{V}_r(A)$ is* irreducible *over $\mathbb{C}$ provided $n > t_{m-r}$.*

Before proving this lemma we first explain the reason for the condition $n > t_{m-r}$ (which is stronger than the condition imposed by the Pataki bound (10)). The variety $\mathcal{V}_r(A)$ is the dual of the determinantal variety $\{x \in \mathbb{C}^n : \operatorname{rank}(A_0 + x_1 A_1 + \cdots + x_n A_n) \leq r\}$. The condition $n > t_{m-r}$ rules out the case where this determinantal variety is zero-dimensional, in which case the dual variety $\mathcal{V}_r(A)$ is a union of hyperplanes and is thus not irreducible. Note that if we are only interested in irreducibility statements over $\mathbb{Q}$ (assuming that $A_0, \ldots, A_n$ are generic with entries in $\mathbb{Q}$) then we do not need to impose such a condition. See [SS15, Remark 2.2] for more on this.

Proof of Lemma 3. The main ingredient of the proof is Bertini's irreducibility theorem [Deb99, Theorem 4.23]. We will start by showing that the variety

$$X = \mathcal{A}(x) + A_0, \ \ XZ = 0, \ \ \operatorname{rank}(X) \leq r, \ \ \operatorname{rank}(Z) \leq m - r \tag{11}$$

is irreducible for a generic choice of $A_0, \ldots, A_n$. In [NRS10, Lemma 6] it was shown that $\{XZ = 0\}^r := \{(X, Z) : XZ = 0, \operatorname{rank}(X) \leq r, \operatorname{rank}(Z) \leq m - r\}$ is irreducible. Consider the projection map $u(X, Z) = X$. We know that $u(\{XZ = 0\}^r)$ is the determinantal variety consisting of symmetric matrices of rank $\leq r$ and has dimension $t_m - t_{m-r}$. By Bertini theorem [Deb99, Theorem 4.23] we know that for a generic affine subspace L of dimension n the variety $u^{-1}(L)$ is going to be irreducible provided $t_m - t_{m-r} \geq 1 + \operatorname{codim} L = 1 + t_m - n$, i.e., provided that $n \geq t_{m-r} + 1$. In other words this tells us that (11) is irreducible for a generic choice of $A_0, \ldots, A_n$.

Consider now the map $\phi(x, X, Z) = (x, X, -Z/(\mathcal{A}^*(Z)^T x), \mathcal{A}^*(Z)/(\mathcal{A}^*(Z)^T x))$ (where the last coordinates stand for c). Observe that the image of the restriction of ϕ to the solution set of (11) is exactly the variety defined by (8). Since ϕ is rational at all points, it is regular [Sha77, Thm 4, Sec. 3.2]. Because the solution set of (11) is irreducible, its image by ϕ is irreducible. Since $\mathcal{V}_r(A)$ is the projection of an irreducible variety it is also irreducible. □

The degrees of the irreducible components $\mathcal{V}_r(A)$ were computed (for generic $A = (A_0, \ldots, A_n)$) in [NRS10, vBR09] and are denoted by $\delta(n, m, r)$. The resulting formulas involve minors of the matrix of binomial coefficients. An elementary analysis of these formulas allows us to show that in a special regime for n and r, the algebraic degree is at least $2^{m^2/20}$.

Lemma 4. *Assume m even and large enough and consider $n = t_{m/2} + 1$ and $r = m/2 + 1$. Then for generic $A = (A_0, \ldots, A_n) \in (\mathbf{S}^m(\mathbb{C}))^{n+1}$ the variety $\mathcal{V}_r(A)$ has degree $\geq 2^{m^2/20}$.*

Proof. The proof is in Appendix A. □

In order to use Lemma 4 we need to show that there is at least one choice of $A = (A_0, \ldots, A_n)$ with $n = t_{m/2} + 1$ such that the variety $\mathcal{V}_r(A)$ with $r = m/2 + 1$ will actually belong to $\partial_a S^o$, where $S = \{x : A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0\}$. We can prove this by appealing to results by Amelunxen and Bürgisser [AB15] where random semidefinite programs were analyzed and where it was shown that every value of rank in the Pataki range occurs with "positive probability".

Lemma 5. *Let m and $1 \leq n \leq t_m$ be fixed. Let r in the associated Pataki range (10) with the additional constraint $n > t_{m-r}$. Let Γ be any Zariski open set in $(\mathbf{S}^m(\mathbb{C}))^{n+1}$. Then there exists $A = (A_0, \ldots, A_n) \in \Gamma \cap (\mathbf{S}^m(\mathbb{R}))^{n+1}$ such that the variety $\mathcal{V}_r(A)$ is contained in $\partial_a S^o$ where $S = \{x \in \mathbb{R}^n : A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0\}$.*

Proof. See Appendix B. □

The proof of Theorem 2 is now complete:

Proof of Theorem 2. Let m be even and large enough and let $n = t_{m/2} + 1$. Lemma 5 with $r = m/2 + 1$ tells us that there is a spectrahedron such that the variety $\mathcal{V}_r(A)$ is contained in $\partial_a S^o$ where $S = \{x : A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0\}$. By Lemma 4 we know that the degree of this variety is at least $2^{m^2/20}$ and so $d = \deg(\partial_a C^o) \geq 2^{m^2/20}$. But this means that $\operatorname{rank}_{\mathrm{psd}}(S) \leq m \leq \sqrt{20 \log d}$ as desired. □

4 Example

In this section we consider an example of spectrahedral shadow to illustrate some of the ideas presented in the proofs of Theorem 1 and Theorem 2.

Consider the following linear matrix inequality:

$$A(x, y, s, t) := \begin{bmatrix} 1+s & t & x+s & y-t \\ t & 1-s & -y-t & x-s \\ x+s & -y-t & 1+x & -y \\ y-t & x-s & -y & 1-x \end{bmatrix}.$$

One can show that the projection of the associated spectrahedron on the variables (x, y) is the regular pentagon in the plane, i.e., if we let S be the spectrahedron associated to A and $\pi(x, y, s, t) = (x, y)$ then:

$$C := \pi(S) = \operatorname{conv}\left\{ \left(\cos\left(\frac{2k\pi}{5}\right), \sin\left(\frac{2k\pi}{5}\right) \right), k = 0, \ldots, 4 \right\}. \tag{12}$$

It is not difficult to see that the polar of C is another regular pentagon but slightly rotated and scaled:

$$C^o = \frac{1}{\cos(\pi/5)} \operatorname{conv}\left\{ \left(\cos\left(\frac{2(k+1/2)\pi}{5}\right), \sin\left(\frac{2(k+1/2)\pi}{5}\right) \right), k = 0, \ldots, 4 \right\}.$$

From Section 2 we know that the KKT equations allow us to get a polynomial that vanishes on the boundary of C^o. The associated variety (denoted by $\mathcal{V}$ in (6)) in this case is shown in Figure 2. We see that the variety $\mathcal{V}$ contains the algebraic boundary of the polar C^o (red lines). However we also see that it has extra components that are not in $\partial_a C^o$: these extra components are shown in blue in Figure 2.

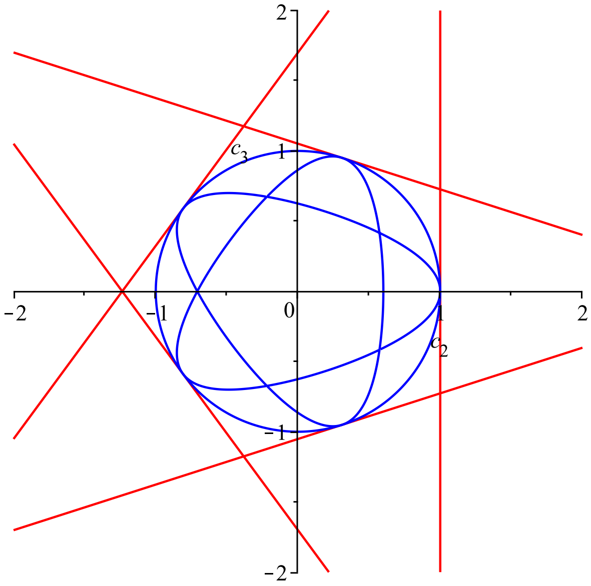

Figure 2: Variety $\mathcal{V}$ defined in (6) that vanishes on the boundary of C^o, where C is the regular pentagon, see (12). We see that $\partial_a C^o \subset \mathcal{V}$ and that $\mathcal{V}$ has extra components not in $\partial_a C^o$. These are shown in blue.

5 Discussion

The algebraic argument given in this paper can also be used to lower bound the size of second-order cone lifts, or more generally lifts using products of $\mathbf{S}_+^k$. More precisely one can show that if C has a lift using r copies of $\mathbf{S}_+^k$ then $r \geq \frac{1}{k^2} \log d$ where d is the degree of the algebraic boundary of C^o. In particular we recover the result of Goemans (Theorem 3) with $k = 1$.

There are a couple of questions that we believe it would be interesting to pursue further:

- **Polytopes:** One important question is to know whether the lower bound $\operatorname{rank}_{\text{psd}}(C) \geq \sqrt{\log d}$ can be improved in the case where C is a polytope? In particular can the lower bound be improved to $\log d$ in this special case? Recall that if C is a polytope then $d = \deg \partial_a C^o$ is simply its number of vertices.

- **Vertices of spectrahedra:** A related question is to know whether the bound of 2^{m^2} on the number of vertices of spectrahedral shadows (Theorem 4) is tight? In words, can we find a spectrahedron (or a spectrahedral shadow) that has $2^{\Omega(m^2)}$ vertices? We believe that a natural candidate to try are random spectrahedra of appropriate dimension $n \approx t_{m/2}$. Results in [AB15] can be useful for this question.

- **Explicit example:** Thirdly, is there an *explicit* example of a spectrahedron whose polar has an algebraic boundary of degree $2^{\Omega(m^2)}$?

- **Analysis of algebraic degree:** Finally we believe it would be useful to have an (asymptotic) analysis of the formulas for the algebraic degrees of semidefinite programming $\delta(n, m, r)$. For this paper we have used elementary manipulations to show that when $n \approx t_{m/2}$ and for certain values of r then $\delta(n, m, r)$ is $2^{\Omega(m^2)}$, but we believe a more complete and systematic study of these quantities can be undertaken. For example we conjecture that in the regime $n \approx t_{m/2}$ the value of $\delta(n, m, r)$ for any r in the Pataki range is $2^{\Omega(m^2)}$. Proving such a result would allow us to simplify the proof of Theorem 2 by bypassing the need for Lemma 5 (it suffices to take any generic spectrahedron of dimension say $n = t_{m/2} + 1$ and to observe that *at least one* of the $\mathcal{V}_r(A)$ must belong to $\partial_a S^o$). An analysis of the values of $\delta(n, m, r)$ would also allow us to improve the constants in Theorems 1 and 2. For example, where we used the Bézout bound in Lemma 1 one can instead use the quantity $\sum_r \delta(n, m, r)$ (where r ranges over the Pataki range) as an upper bound on the number of solutions of the KKT system.

Acknowledgments. We would like to thank Rainer Sinn for clarifications concerning questions of irreducibility used in Lemma 3, and James Saunderson for comments that helped improve the paper. We would also like to thank anonymous referees for their comments on an earlier version of the paper that was submitted for presentation to the conference MEGA 2017. Finally we would like to thank Xavier Allamigeon and Stéphane Gaubert for organizing the session on Semidefinite Optimization at the PGMO 2016 conference where this project was started. The second author is supported by the ANR JCJC GALOP grant.

A Proof of Lemma 4: analysis of the formula for the algebraic degree of semidefinite programming

In this subsection we prove Lemma 4 which we restate below.

Lemma (Restatement of Lemma 4). *Assume m even and large enough and consider $n = t_{m/2} + 1$ and $r = m/2 + 1$. Then for generic $A = (A_0, \ldots, A_n) \in (\mathbf{S}^m(\mathbb{C}))^{n+1}$ the variety $\mathcal{V}_r(A)$ has degree $\geq 2^{m^2/20}$.*

For this we rely on the formula for the algebraic degree of semidefinite programming proved in [vBR09].

Let $\delta(n, m, r)$ be the degree of the variety $\mathcal{V}_r(A)$ where A is a generic pencil $(A_0, \ldots, A_n) \in (\mathbf{S}^m(\mathbb{R}))^{n+1}$. A formula for $\delta(n, m, r)$ was given in [vBR09] which we describe now. Let Ψ be the (infinite) matrix of binomial coefficients, i.e., $\Psi_{i,j} = \binom{i}{j}$ for $i, j \geq 0$. For $I \subseteq \{1, \ldots, m\}$ define ψ_I to be the sum of all the $|I| \times |I|$ minors of $\Psi[I, \cdot]$. For example if I is a singleton we have $\psi_{\{i\}} = 2^{i-1}$.

Theorem 5 ([vBR09], see also [Ran12]). *For a generic $A = (A_0, \ldots, A_n)$ the algebraic degree of $\mathcal{V}_r(A)$ (see Equation (9)) is given by:*

$$\delta(n, m, r) = \sum_{\substack{I \subseteq \{1, \ldots, m\} \\ |I| = m - r, s(I) = n}} \psi_I \psi_{I^c} \tag{13}$$

where for $I \subseteq \{1, \ldots, m\}$ we denote by $s(I)$ the sum of the elements of I, and $I^c = \{1, \ldots, m\} \setminus I$.

The main purpose of this Appendix is to prove the following lower bound on $\delta(n, m, r)$ in a special regime.

Lemma 6. *For all large enough even m, $n = t_{m/2}+1$ and $r = m/2+1$ we have $\delta(n,m,r) \geq 2^{m^2/20}$.*

The bounds we give in this appendix are very crude and are not meant to be optimal. We actually conjecture that in the regime $n \approx t_{m/2}$, we have $\delta(t_{m/2}, m, r) \geq 2^{\Omega(m^2)}$ for any r in the Pataki range (10).

In order to prove our result we will first analyze the value of ψ on intervals. We will show

Lemma 7. *For any integers $p \leq q$ we have $\psi_{[p+1,q]} \geq (1 + \frac{q-p}{2p-1})^{t_p}$.*

Before proving Lemma 7, we first see how to use it to prove Lemma 6.

Proof of Lemma 6. Consider $I = \{1, \dots, m/2-2\} \cup \{m\}$. Then $|I| = m/2-1 = m-r$ and $s(I) = t_{m/2}+1 = n$. Thus $\delta(n,m,r) \geq \psi_I \psi_{I^c} \geq \psi_{I^c} = \psi_{[m/2-2,m-1]}$. Using Lemma 7 we get

$$\psi_{[m/2-2,m-1]} \geq \left(1 + \frac{m/2+1}{m-3}\right)^{t_{m/2-2}}.$$

We now use the fact that $1 + \frac{m/2+1}{m-3} \geq 1 + 1/2 = 3/2$ and $t_{m/2-2} \geq m^2/9$ for large enough m to get $\psi_{[m/2-1,m-2]} \geq 2^{(\log_2(3/2)/9)m^2} \geq 2^{m^2/20}$. □

It thus remains to prove Lemma 7. We can get the value of ψ on intervals by considering the case $n = t_{m-r}$ in (13). Indeed in this case there is only one set I that satisfies the constraints of the summation (13) which is $I = \{1, \dots, m-r\}$. Since $\psi_{[1,m-r]} = 1$ it follows that $\delta(t_{m-r}, m, r) = \psi_{[m-r+1,m]}$. On the other hand a simpler formula for $\delta(t_{m-r}, m, r)$ was provided in [NRS10, Corollary 15], based on a result by Harris and Tu [HT84]. This tells us that

$$\delta(t_{m-r}, m, r) = \psi_{[m-r+1,m]} = \prod_{i=0}^{m-r-1} \frac{\binom{m+i}{m-r-i}}{\binom{2i+1}{i}}. \tag{14}$$

The formula on the right-hand side can be simplified further using simple manipulations to get

$$\psi_{[m-r+1,m]} = \prod_{0 \leq i \leq j \leq m-r-1} \frac{r+i+j+1}{i+j+1}. \tag{15}$$

To see why this holds, first use the definition of binomial coefficient $\binom{n}{k} = \frac{n \dots (n-k+1)}{k!}$ to get

$$\psi_{[m-r+1,m]} = \prod_{i=0}^{m-r-1} \frac{\binom{m+i}{m-r-i}}{\binom{2i+1}{i}} = \prod_{i=0}^{r-1} \frac{(m+i)\dots(r+2i+1)}{(m-r-i)!} \cdot \frac{i!}{(2i+1)\dots(i+2)} \tag{16}$$

Separating the terms in (16) we get

$$\psi_{[m-r+1,m]} = \left[\prod_{0 \leq i \leq j \leq m-r-1} (r+i+j+1)\right] \cdot \left[\prod_{i=0}^{m-r-1} \frac{i!}{(m-r-i)!(2i+1)\dots(i+2)}\right]. \tag{17}$$

Noting that $\prod_{i=0}^{m-r-1} i!/(m-r-i)! = \frac{1}{(m-r)!} = \prod_{i=0}^{m-r-1} \frac{1}{(i+1)}$ we see that the second factor in (17) is equal to

$$\prod_{i=0}^{m-r-1} \frac{1}{(2i+1)\dots(i+2)(i+1)} = \prod_{i=0}^{m-r-1} \prod_{j=0}^{i} \frac{1}{i+j+1}. \tag{18}$$

By doing an appropriate change of variables and plugging this back in (17) we get (15).

Now to prove the bound of Lemma 7 note that each term in the product (15) is at least $1 + \frac{r}{2(m-r)-1}$ and that there are t_{m-r} terms in the product. The statement of Lemma 7 corresponds to $p = m-r$ and $q = m$. This completes the proof.

B Proof of Lemma 5: occurrence of each value of rank in the Pataki range

In this Appendix we prove Lemma 5 which we restate here for convenience.

Lemma (Restatement of Lemma 5)**.** *Let m and $1 \leq n \leq t_m$ be fixed. Let r in the associated Pataki range (10) with the additional constraint $n > t_{m-r}$. Let Γ be any Zariski open set in $(\mathbf{S}^m(\mathbb{C}))^{n+1}$. Then there exists a pencil $A = (A_0, \dots, A_n) \in \Gamma \cap (\mathbf{S}^m(\mathbb{R}))^{n+1}$ such that the variety $\mathcal{V}_r(A)$ is contained in $\partial_a S^o$, where $S = \{x \in \mathbb{R}^n : A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0\}$.*

Proof. For convenience in this proof we let, for $A = (A_0, \dots, A_n) \in (\mathbf{S}^m(\mathbb{R}))^{n+1}$, $S(A) \subset \mathbb{R}^n$ denote the associated spectrahedron:

$$S(A) = \{x \in \mathbb{R}^n : A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0\}.$$

In the paper [AB15, Remark 4.1] it is shown that for any r satisfying the Pataki bounds (10) we have

$$\Pr_{A_0,\dots,A_n,c}\left[\operatorname{rank}\left(\operatorname*{argmax}_{x\in S(A)} c^T x\right) = r\right] > 0 \tag{19}$$

where $A_0, \dots, A_n, c$ are standard Gaussian with respect to the Euclidean inner product. In other words, each value in the Pataki range occurs with positive probability. Fix r in the Pataki range satisfying $n > t_{m-r}$ and consider

$$\Omega_r = \left\{(A_0, \dots, A_n) \in (\mathbf{S}^m(\mathbb{R}))^{n+1} : \Pr_c\left[\operatorname{rank}\left(\operatorname*{argmax}_{x\in S(A)} c^T x\right) = r\right] > 0\right\}.$$

By (19) we know that Ω_r has positive probability (otherwise the complement of Ω_r has probability 1 which would contradict (19)). Thus this means that Ω_r must meet Γ since Γ is Zariski open.

Let $A := (A_0, \dots, A_n) \in \Omega_r \cap \Gamma$ and let $S = S(A) = \{x \in \mathbb{R}^n : A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0\}$. To prove our claim we will show that $\mathcal{V}_r(A)$ intersects the boundary ∂S^o along a semialgebraic set of dimension $n-1$. This will prove our claim because if we let U be this semialgebraic set we then have on the one hand $\partial_a S^o \supseteq \bar{U}^Z$ (where $\bar{U}^Z$ denotes the Zariski closure) and on the other hand $\bar{U}^Z = \mathcal{V}_r(A)$, the latter following from the fact that $\mathcal{V}_r(A)$ is irreducible of dimension $n-1$ and that $\dim_{\mathbb{C}}(\bar{U}^Z) = n-1$ since U is a semialgebraic set of dimension $n-1$, see [BCR13, Proposition 2.8.2].

It remains to show that $\mathcal{V}_r(A)$ intersects ∂S^o along a semialgebraic set of dimension $n-1$. To see why this holds let

$$U = \tilde{U} \cap \partial S^o \quad \text{where} \quad \tilde{U} = \left\{c \in \mathbb{R}^n : \operatorname{rank}\left(\operatorname*{argmax}_{x\in S} c^T x\right) = r\right\}.$$

By definition of $\mathcal{V}_r(A)$ (recall that $\mathcal{V}_r(A)$ is defined in terms of rank-constrained KKT equations) we have $U \subseteq \mathcal{V}_r(A) \cap \partial S^o$. Now observe that U is a semialgebraic set of dimension $n-1$: indeed note that $\tilde{U}$ has nonempty interior (since it is a semialgebraic set with positive probability, see Lemma 8) and so $U = \tilde{U} \cap \partial S^o$ has dimension $n-1$ since for any $\alpha \in \partial S^o$ and neighborhood A of α, $\dim(A \cap \partial S^o) = n-1$ (because ∂S^o is the boundary of a full-dimensional convex set). This completes the proof. □

Lemma 8. *If $W \subseteq \mathbb{R}^N$ is semialgebraic and W has positive probability under the standard Gaussian measure, then W has nonempty interior.*

Proof. Any semialgebraic set can be decomposed as a disjoint union of semialgebraic sets that are homeomorphic to $(0,1)^d$ (see [BCR13, Theorem 2.3.6]). Since $\Pr[W] > 0$, W must have a component that is homeomorphic to $(0,1)^N$ and thus W has nonempty interior. □

References

[AB15] Dennis Amelunxen and Peter Bürgisser. Intrinsic volumes of symmetric cones and applications in convex programming. *Mathematical Programming*, 149(1-2):105–130, 2015. 8, 9, 12

[BCR13] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real algebraic geometry*, volume 36. Springer Science & Business Media, 2013. 12, 13

[BTN01] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, 2001. 3

[dCST15] Marcel K. de Carli Silva and Levent Tunçel. Vertices of spectrahedra arising from the elliptope, the theta body, and their relatives. *SIAM Journal on Optimization*, 25(1):295–316, 2015. 6

[Deb99] Olivier Debarre. Introduction à la géométrie algébrique. 1999. Available online at `http://www.math.ens.fr/~debarre/DEA99.pdf`. 7

[FGP$^+$15] Hamza Fawzi, João Gouveia, Pablo A. Parrilo, Richard Z. Robinson, and Rekha R. Thomas. Positive semidefinite rank. *Mathematical Programming*, 153(1):133–177, 2015. 1

[Goe15] Michel X. Goemans. Smallest compact formulation for the permutahedron. *Mathematical Programming*, 153(1):5–11, 2015. 2

[GPT13] João Gouveia, Pablo A. Parrilo, and Rekha R. Thomas. Lifts of convex sets and cone factorizations. *Mathematics of Operations Research*, 38(2):248–264, 2013. 1, 2

[HT84] Joe Harris and Loring W. Tu. On symmetric and skew-symmetric determinantal varieties. *Topology*, 23(1):71–84, 1984. 11

[LP95] Monique Laurent and Svatopluk Poljak. On a positive semidefinite relaxation of the cut polytope. *Linear Algebra and its Applications*, 223:439–461, 1995. 6

[NRS10] Jiawang Nie, Kristian Ranestad, and Bernd Sturmfels. The algebraic degree of semidefinite programming. *Mathematical Programming*, 122(2):379–405, 2010. 2, 4, 7, 8, 11

[Ran12] Kristian Ranestad. Algebraic degree in semidefinite and polynomial optimization. In *Handbook on Semidefinite, Conic and Polynomial Optimization*, pages 61–75. Springer, 2012. 10

[Roc97] R Tyrell Rockafellar. *Convex analysis*, volume 28. Princeton University Press, 1997. 2

[RS12] Philipp Rostalski and Bernd Sturmfels. Dualities in convex algebraic geometry. In Grigoriy Blekherman, Pablo A. Parrilo, and Rekha R. Thomas, editors, *Semidefinite Optimization and Convex Algebraic Geometry*, chapter 5, pages 203–249. SIAM, 2012. 2

[Sch13] Rolf Schneider. *Convex bodies: the Brunn–Minkowski theory*. Number 151. Cambridge University Press, 2013. 6

[Sha77] Igor Shafarevich. *Basic Algebraic Geometry 1*. Springer Verlag, 1977. 7

[Sin15] Rainer Sinn. Algebraic boundaries of convex semi-algebraic sets. *Research in the Mathematical Sciences*, 2(1):1, 2015. 5

[SS15] Rainer Sinn and Bernd Sturmfels. Generic spectrahedral shadows. *SIAM Journal on Optimization*, 25(2):1209–1220, 2015. 2, 7

[vBR09] Hans-Christian Graf von Bothmer and Kristian Ranestad. A general formula for the algebraic degree in semidefinite programming. *Bulletin of the London Mathematical Society*, 41(2):193–197, 2009. 2, 8, 10