MATRIX CUBES PARAMETERIZED BY EIGENVALUES

JIAWANG NIE* AND BERND STURMFELS†

Abstract. An elimination problem in semidefinite programming is solved by means of tensor algebra. It concerns families of matrix cube problems whose constraints are the minimum and maximum eigenvalue function on an affine space of symmetric matrices. An LMI representation is given for the convex set of all feasible instances, and its boundary is studied from the perspective of algebraic geometry. This generalizes the known LMI representations of k-ellipses and k-ellipsoids.

Key words. Linear matrix inequality (LMI), semidefinite programming (SDP), matrix cube, tensor product, tensor sum, k-ellipse, algebraic degree

AMS subject classifications. 14P10, 34K20, 65K10, 90C22

1. Introduction. The *matrix cube problem* in semidefinite programming is concerned with the following question. Given real symmetric $N\times N$-matrices $A_0, A_1, \ldots, A_m$, does every point $(t_1,\ldots,t_m)$ in the cube $\prod_{i=1}^m [\lambda_i, \mu_i]$ satisfy the matrix inequality

$$A_0 + \sum_{k=1}^m t_k A_k \succeq 0 \ ?$$

The inequality means that the symmetric matrix $A_0+\sum_{k=1}^m t_k A_k$ is positive semidefinite, i.e. its N eigenvalues are all non-negative reals. This problem is NP-hard [3], and it has important applications in robust optimization and control [4, 11], e.g., in Lyapunov stability analysis for uncertain dynamical systems, and for various combinatorial problems which can be reduced to maximizing a positive definite quadratic form over the unit cube. For a recent study see [5].

In this paper we examine parameterized families of matrix cube problems, where the lower and upper bounds that specify the cube are the eigenvalues of symmetric matrices that range over a linear space of matrices. We define the set

$$\mathcal{C} \ = \ \left\{ (x,d) \in \mathbb{R}^n \times \mathbb{R} \ \middle| \ \begin{array}{r} d\cdot A_0 + \sum_{k=1}^m t_k A_k \succeq 0 \ \text{ whenever} \\ \lambda_{\min}(B_k(x)) \le t_k \le \lambda_{\max}(B_k(x)) \\ \text{for } \ k=1,2,\ldots,m \end{array} \right\}, \tag{1.1}$$

where the A_i are constant symmetric matrices of size $N_0 \times N_0$, the symbols $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the minimum and maximum eigenvalues of a matrix, and the $N_k \times N_k$ matrices $B_k(x)$ are linear matrix polynomials of the form

$$B_k(x) \ = \ B_0^{(k)} + x_1 B_1^{(k)} + \cdots + x_n B_n^{(k)}.$$

Here $B_0^{(k)}, B_1^{(k)}, \ldots, B_n^{(k)}$ are constant symmetric $N_k \times N_k$ matrices for all k. Note that the standard matrix cube is a special case of the set (1.1). For instance, if every $B_k(x)$ is the following 2×2 constant symmetric matrix

$$\frac{1}{2}\begin{bmatrix} \lambda_i+\mu_i & \lambda_i-\mu_i \\ \lambda_i-\mu_i & \lambda_i+\mu_i \end{bmatrix},$$

* Department of Mathematics, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA (njw@math.ucsd.edu).

† Department of Mathematics, University of California, Berkeley, CA 94720, USA (bernd@math.berkeley.edu). Research was supported by an Alexander von Humboldt research prize and NSF grant DMS-0456960.

then the set defined by (1.1) is a standard matrix cube.

Since the minimum eigenvalue function $\lambda_{\min}(\cdot)$ is concave and the maximum eigenvalue function $\lambda_{\max}(\cdot)$ is convex, it follows that $\mathcal{C}$ is a convex subset in $\mathbb{R}^{n+1}$. To see this, we suppose that $(x^{(1)}, d^{(1)})$ and $(x^{(2)}, d^{(2)})$ are in $\mathcal{C}$. The concavity of $\lambda_{\min}(\cdot)$ and the convexity of $\lambda_{\max}(\cdot)$ imply that, for each index k,

$$\frac{1}{2}\left(\lambda_{min}(B_k(x^{(1)}) + \lambda_{min}(B_k(x^{(2)})\right) \leq \lambda_{min}\left(B_k\left(\frac{x^{(1)}+x^{(2)}}{2}\right)\right),$$
$$\frac{1}{2}\left(\lambda_{max}(B_k(x^{(1)}) + \lambda_{max}(B_k(x^{(2)})\right) \geq \lambda_{max}\left(B_k\left(\frac{x^{(1)}+x^{(2)}}{2}\right)\right).$$

Therefore, the cube

$$\prod_{k=1}^{m}\left[\lambda_{min}\left(B_k\left(\frac{x^{(1)}+x^{(2)}}{2}\right)\right),\ \lambda_{max}\left(B_k\left(\frac{x^{(1)}+x^{(2)}}{2}\right)\right)\right]$$

is contained in the intersection

$$\prod_{k=1}^{m}\left[\lambda_{min}(B_k(x^{(1)})), \lambda_{max}(B_k(x^{(1)}))\right] \bigcap \prod_{k=1}^{m}\left[\lambda_{min}(B_k(x^{(2)})), \lambda_{max}(B_k(x^{(2)}))\right].$$

So, we must have $\frac{1}{2}\{(x^{(1)}, d^{(1)}) + (x^{(2)}, d^{(2)})\} \in \mathcal{C}$, that is, $\mathcal{C}$ is convex.

Our definition of $\mathcal{C}$ involves a polynomial system in free variables (x, d) and universally quantified variables t. Quantifier elimination in real algebraic geometry [1] tells us that $\mathcal{C}$ is semialgebraic, which means that it can be described by a Boolean combination of polynomial equalities or inequalities in (x, d). However, to compute such a description by algebraic elimination algorithms is infeasible.

Linear matrix inequalities (LMI) are a useful and efficient tool in system and control theory [2]. LMI representations are convenient for building convex optimization models, especially in semidefinite programming [16]. The set described by an LMI is always convex and semialgebraic. Since our set $\mathcal{C}$ is convex and semialgebraic, it is natural to ask whether $\mathcal{C}$ admits an LMI representation.

Our aim is to answer this question affirmatively. Theorem 3.2 states that

$$\mathcal{C} \;=\; \{\,(x,d) \in \mathbb{R}^n \times \mathbb{R} :\, \mathcal{L}(x,d) \succeq 0\,\}, \tag{1.2}$$

where $\mathcal{L}(x, d)$ is a linear matrix polynomial whose coefficients are larger symmetric matrices that are constructed from the matrices A_i and $B_j^{(k)}$. The construction involves operations from tensor algebra and is carried out in Section 3.

First, however, some motivation is needed. In Section 2 we shall explain why the convex semialgebraic sets $\mathcal{C}$ are interesting. This is done by discussing several geometric applications, notably the study of m-ellipses and m-ellipsoids [12].

Section 4 is devoted to algebraic geometry questions related to our set $\mathcal{C}$: What is the Zariski closure $\mathcal{Z}(\partial\mathcal{C})$ of the boundary $\partial\mathcal{C}$? What is the polynomial defining $\mathcal{Z}(\partial\mathcal{C})$? What is the degree of $\mathcal{Z}(\partial\mathcal{C})$? Is the boundary $\partial\mathcal{C}$ irreducible ?

In Section 5, we discuss an application to robust control, explore the notion of matrix ellipsoids, and conclude with some directions for further research.

2. Ellipses and beyond. In this section we illustrate the construction of the set $\mathcal{C}$ for some special cases. The first case to consider is $N_1 = \cdots = N_m = 1$ when each $B_k(x) = b_k^T x + \beta_k$ is a linear scalar function. Then our elimination problem is solved as follows:

$$\mathcal{C} \;=\; \left\{ (x,d) \in \mathbb{R}^{n+1} :\, d \cdot A_0 + \sum_{k=1}^{m} (b_k^T x + \beta_k) A_k \succeq 0 \right\}.$$

Thus $\mathcal{C}$ is a spectrahedron (the solution set of an LMI) and, conversely every spectrahedron arises in this way. The Zariski closure $\mathcal{Z}(\partial\mathcal{C})$ of its boundary $\partial\mathcal{C}$ is the hypersurface defined by the vanishing of the determinant of the above matrix. For generic data A_i, b_k, β_k this hypersurface is irreducible and has degree N_0. Throughout this paper, we shall use the term "generic" in the sense of algebraic geometry (cf. [13]). Randomly chosen data are generic with probability one.

A natural extension of the previous example is the case when each given matrix $B_k(x)$ is diagonal. We write this diagonal matrix as follows:

$$B_k(x) \;=\; \operatorname{diag}\left({b_k^{(1)}}^T x + \beta_k^{(1)}, \ldots, {b_k^{(N_k)}}^T x + \beta_k^{(N_k)} \right)$$

Then our spectrahedron can be described by an intersection of LMIs:

$$\mathcal{C} \;=\; \left\{ (x,d) :\, d \cdot A_0 + \sum_{k=1}^{m} ({b_k^{(i_k)}}^T x + \beta_k^{(i_k)}) A_k \succeq 0,\;\; 1 \le i_k \le N_k, 1 \le k \le m \right\}.$$

The Zariski closure $\mathcal{Z}(\partial\mathcal{C})$ of the boundary $\partial\mathcal{C}$ is a hypersurface which is typically reducible. It is defined by the product of all determinants of the above matrices which contribute an active constraint for $\mathcal{C}$. Each of the determinants has degree N_0 and there can be as many as $N_1 N_2 \cdots N_m$ of these boundary components.

The point of departure for this project was our paper with Parrilo [12] whose result we briefly review. Given points $(u_1, v_1), \ldots, (u_m, v_m)$ in the plane $\mathbb{R}^2$ and a parameter $d > 0$, the corresponding *m-ellipse* $\mathcal{E}_d$ is the set of all points (x_1, x_2) whose sum of distances to the given points (u_i, v_i) is at most d. In symbols,

$$\mathcal{E}_d \;=\; \left\{ x \in \mathbb{R}^2 \,:\, \sum_{k=1}^{m} \sqrt{(x_1 - u_k)^2 + (x_2 - v_k)^2} \,\le\, d \right\}.$$

In [12] it is shown that $\mathcal{E}_d$ is a spectrahedron, an explicit LMI representation of size $2^m \times 2^m$ is given, and the degree of $\partial\mathcal{E}$ is shown to be 2^m when m is odd, and $2^m - \binom{m}{m/2}$ when m is even. For instance, if $m = 3$, $(u_1, v_1) = (0,0), (u_2, v_2) = (1,0)$ and $(u_3, v_3) = (0,1)$, then the 3-ellipse $\mathcal{E}_d$ has the LMI representation

$$\begin{bmatrix}
d+3x_1-1 & x_2-1 & x_2 & 0 & x_2 & 0 & 0 & 0 \\
x_2-1 & d+x_1-1 & 0 & x_2 & 0 & x_2 & 0 & 0 \\
x_2 & 0 & d+x_1+1 & x_2-1 & 0 & 0 & x_2 & 0 \\
0 & x_2 & x_2-1 & d-x_1+1 & 0 & 0 & 0 & x_2 \\
x_2 & 0 & 0 & 0 & d+x_1-1 & x_2-1 & x_2 & 0 \\
0 & x_2 & 0 & 0 & x_2-1 & d-x_1-1 & 0 & x_2 \\
0 & 0 & x_2 & 0 & x_2 & 0 & d-x_1+1 & x_2-1 \\
0 & 0 & 0 & x_2 & 0 & x_2 & x_2-1 & d-3x_1+1
\end{bmatrix} \succeq 0.$$

See Figure 2.1 for a picture of this 3-ellipse and the Zariski closure of its boundary. Here, the three marked points are the foci, the innermost curve surrounding the three foci is the 3-ellipse, and the three outer curves appear in its Zariski closure.

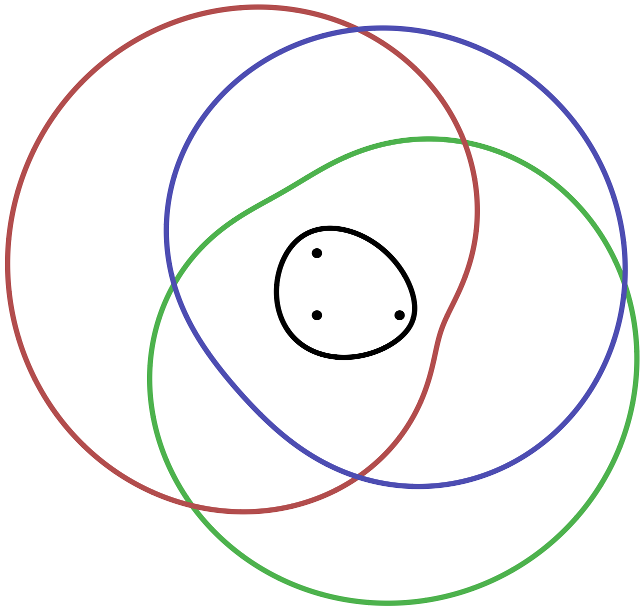

FIG. 2.1. *The Zariski closure of the 3-ellipse is a curve of degree eight.*

We can model the m-ellipse using matrix cubes as in (1.1) as follows. Let $N_0 = 1$ and define each scalar A_i to be 1, and let $N_1 = \cdots = N_m = 2$ and define

$$B_k(x) \;=\; \begin{bmatrix} x_1 - u_k & x_2 - v_k \\ x_2 - v_k & u_k - x_1 \end{bmatrix} \quad \text{for } \; k = 1, \cdots, m.$$

The eigenvalues of $B_k(x)$ are $\pm\sqrt{(x_1 - u_k)^2 + (x_2 - v_k)^2}$, and we find that

$$\mathcal{C} = \big\{(x, d) \in \mathbb{R}^3 \,:\, d \geq t_1 + \cdots + t_m \;\; \text{whenever } |t_k| \leq \sqrt{(x_1 - u_k)^2 + (x_2 - v_k)^2}\big\}.$$

This formula characterizes the parametric m-ellipse

$$\mathcal{C} \;=\; \big\{(x_1, x_2, d) \in \mathbb{R}^3 \;:\; (x_1, x_2) \in \mathcal{E}_d\big\},$$

and its LMI representation (1.2) is precisely that given in [12, Theorem 2.3]. The construction extends to m-ellipsoids, where the points lie in a higher-dimensional space. We shall return to this topic and its algebraic subtleties in Example 4.4. In Section 5 we introduce a matrix version of the m-ellipses and m-ellipsoids.

Consider now the case when $N_0 = 1$ and $A_0 = A_1 = \cdots = A_m = -1$ but the $B_k(x)$ are allowed to be arbitrary symmetric $N_k \times N_k$-matrices whose entries are linear in the unknowns x. Then the spectrahedron in (1.1) has the form

$$\mathcal{C} \;=\; \Big\{(x, d) \in \mathbb{R}^n \times \mathbb{R} \,:\, \sum_{k=1}^{m} \lambda_{\max}(B_k(x)) \leq d\Big\}.$$

A large class of important convex functions on $\mathbb{R}^n$ can be represented in the form $x \mapsto \lambda_{\max}(B_k(x))$. Our main result in the next section gives a recipe for constructing an explicit LMI representation for the graph of a sum of such convex functions. The existence of such a representation is not obvious, given that the Minkowski sums of two spectrahedra is generally not a spectrahedron [14, §3.1].

3. Derivation of the LMI representation. In this section we show that the convex set $\mathcal{C}$ in (1.1) is a spectrahedron, and we apply tensor operations as in [12] to derive an explicit LMI representation. We begin with a short review of tensor sum and tensor product of square matrices. Given two square matrices $F = (F_{ij})_{1\le i,j\le r}$ and $G = (G_{k\ell})_{1\le k,\ell\le s}$, their standard *tensor product* $\otimes$ is the block matrix of format $rs \times rs$ which defined as

$$F \otimes G \;\;=\;\; \big(F_{ij}G\big)_{1\le i,j\le r}.$$

Based on tensor product $\otimes$, we define the *tensor sum* $\oplus$ to be the $rs \times rs$-matrix

$$F \oplus G \;\;=\;\; F \otimes I_s + I_r \otimes G.$$

Here I_r denotes the identity matrix of size $r \times r$. For instance, the tensor sum of two 2×2-matrices is given by the formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \oplus \begin{bmatrix} e & f \\ g & h \end{bmatrix} \;\;=\;\; \begin{bmatrix} a+e & f & b & 0 \\ g & a+h & 0 & b \\ c & 0 & d+e & f \\ 0 & c & g & d+h \end{bmatrix}.$$

Tensor products and tensor sums of matrices are also known as *Kronecker products* and *Kronecker sums* [9]. Note that the tensor operations $\otimes$ and $\oplus$ are not commutative but they are associative. Hence we can remove parentheses when we take the tensor product or tensor sum of k matrices in a fixed order.

The eigenvalues of $F \otimes G$ are the products of pairs of eigenvalues of F and G. Similarly, the eigenvalues of $F \oplus G$ are the sums of such pairs. This is well-known for tensor products, but perhaps slightly less so for tensor sums. We therefore explicitly state the following lemma on diagonalization of tensor sums.

Lemma 3.1. [12, Lemma 2.2] *Let $M_1, \ldots, M_k$ be symmetric matrices, $U_1, \ldots, U_k$ orthogonal matrices, and $\Lambda_1, \ldots, \Lambda_k$ diagonal matrices such that $M_i = U_i \cdot \Lambda_i \cdot U_i^T$ for $i = 1, \ldots, k$. Then the tensor sum transforms as follows:*

$$(U_1 \otimes \cdots \otimes U_k)^T \cdot (M_1 \oplus \cdots \oplus M_k) \cdot (U_1 \otimes \cdots \otimes U_k) \;\;=\;\; \Lambda_1 \oplus \cdots \oplus \Lambda_k.$$

In particular, the eigenvalues of $M_1 \oplus M_2 \oplus \cdots \oplus M_k$ are the sums $\lambda_1 + \lambda_2 + \cdots + \lambda_k$ where λ_1 is any eigenvalue of M_1, λ_2 is any eigenvalue of M_2, etc.

We now turn to the construction of an LMI representation for the convex semi-algebraic set $\mathcal{C}$. First, we introduce a linear operator $\mathcal{A}$ on the space of linear matrix polynomials in $m+1$ unknowns $a = (a_0, a_1, \ldots, a_m)$. Namely, let $P_a(x,d)$ be such a linear matrix polynomial of the form

$$P_a(x,d) \;\;=\;\; a_0 P_0(x,d) + a_1 P_1(x,d) + \cdots + a_m P_m(x,d) \tag{3.1}$$

where the coefficients $P_0(x,d), \ldots, P_m(x,d)$ are matrix polynomials in (x,d). Using the matrices $A_0, A_1, \ldots, A_n$ in (1.1), we define the linear operator $\mathcal{A}$ by

$$\mathcal{A}\big(P_a(x,d)\big) \;:=\; P_0(x,d) \otimes A_0 + P_1(x,d) \otimes A_1 + \cdots + P_m(x,d) \otimes A_m.$$

Second, using the matrices $B_k(x)$ in (1.1) we define the following tensor sum

$$L_a(x,d) \;:=\; (da_0) \oplus (a_1 B_1(x)) \oplus \;\cdots\; \oplus (a_m B_m(x)).$$

This expression is linear matrix polynomial in $a = (a_0, a_1, \dots, a_m)$, and we set

$$\mathcal{L}(x,d) \;:=\; \mathcal{A}\big(L_a(x,d)\big). \tag{3.2}$$

Since $L_a(x,d)$ is linear in both (x,d) and a, and since $\mathcal{A}$ is a linear operator, the matrix $\mathcal{L}(x,d)$ depends linearly on (x,d) and on the matrices $A_0, A_1, \dots, A_m$.

THEOREM 3.2. *The convex semialgebraic set $\mathcal{C} \subset \mathbb{R}^{n+1}$ in (1.1) is a spectrahedron, and it can be represented by the LMI (1.2).*

This theorem implies that optimizing a linear function over the convex set $\mathcal{C}$ is an instance of semidefinite programming [16]. For instance, for a fixed value d^* of d, to minimize a linear functional $c^T x$ over its section $\mathcal{C} \cap \{ (x,d) : d = d^* \}$ is equivalent to solving the semidefinite programming problem

$$\text{minimize } \; c^T x \quad \text{subject to} \quad \mathcal{L}(x,d^*) \succeq 0.$$

The dimension of the matrix $\mathcal{L}(x,d^*)$ is $N_0 N_1 \cdots N_m$. When this number is not too large, the above optimization problem can be solved efficiently by standard SDP solvers like `SeDuMi` [15]. When $N_0 N_1 \cdots N_m$ is big, the above semidefinite programming problem would be quite difficult to solve. Addressing the issue of large-scale SDP lies beyond the scope of this paper. In any case, it would be very useful to find a smaller sized LMI representation for $\mathcal{C}$, but at present we do not know how to reduce the size of the LMI representation of $\mathcal{C}$. We regard this as an interesting research problem, to be revisited in Section 5.

To prove Theorem 3.2 we need one more fact concerning the linear operator $\mathcal{A}$, namely, that $\mathcal{A}$ is invariant under congruence transformations.

LEMMA 3.3. *Let $P_a(x,d)$ be as in (3.1). For any matrix $U(x,d)$ we have*

$$\mathcal{A}\big(U(x,d)^T P_a(x,d) U(x,d)\big) = (U(x,d) \otimes I_{N_0})^T \mathcal{A}\big(P_a(x,d)\big)(U(x,d) \otimes I_{N_0}). \tag{3.3}$$

Proof. First note that tensor product satisfies $(A \otimes B)^T = A^T \otimes B^T$ and

$$(M_1 \otimes M_2) \cdot (M_3 \otimes M_4) \cdot (M_5 \otimes M_6) = (M_1 \cdot M_3 \cdot M_5) \otimes (M_2 \cdot M_4 \cdot M_6)$$

Using these identifies we perform the following direct calculation:

$$\begin{aligned}
&\Big(U(x,d) \otimes I_{N_0}\Big)^T \mathcal{A}\Big(P_a(x,d)\Big) \Big(U(x,d) \otimes I_{N_0}\Big) \\
=\; &\Big(U(x,d) \otimes I_{N_0}\Big)^T \left(\sum_{k=0}^m P_k(x,d) \otimes A_k\right) \Big(U(x,d) \otimes I_{N_0}\Big) \\
=\; &\sum_{k=0}^m \Big(U(x,d) \otimes I_{N_0}\Big)^T \Big(P_k(x,d) \otimes A_k\Big)\Big(U(x,d) \otimes I_{N_0}\Big) \\
=\; &\sum_{k=0}^m \Big(U(x,d)^T P_k(x,d) U(x,d)\Big) \otimes A_k.
\end{aligned}$$

By definition of $\mathcal{A}$, this expression is equal to the left hand side of (3.3). □

Proof. [Proof of Theorem 3.2] Let $U_k(x)$ be orthogonal matrices such that

$$U_k(x)^T \cdot B_k(x) \cdot U_k(x) \;=\; \operatorname{diag}\Big(\lambda_1^{(k)}(x), \dots, \lambda_{N_k}^{(k)}(x)\Big) \;=:\; D_k. \tag{3.4}$$

Here $\lambda_j^{(k)}(x)$ are the algebraic functions representing eigenvalues of the matrices $B_k(k)$. If we set $Q = (1) \otimes U_1(x) \otimes \cdots \otimes U_m(x)$ then, by Lemma 3.1, we have

$$Q^T \cdot L_a(x,d) \cdot Q \;=\; (da_0) \oplus (a_1 D_1(x)) \oplus \cdots \oplus (a_m D_m(x)) \;=:\; \tilde{L}.$$

Note that $\tilde{L}$ is a diagonal matrix, with each diagonal entry having the form

$$d \cdot a_0 \,+\, \sum_{k=1}^{m} \lambda_{j_k}^{(k)}(x) a_k.$$

It follows that $\mathcal{A}(\tilde{L})$ is a block diagonal matrix, with each block of the form

$$d \cdot A_0 \,+\, \sum_{k=1}^{m} \lambda_{j_k}^{(k)}(x) A_k. \tag{3.5}$$

By Lemma 3.3 and the definition of $\mathcal{L}(x,d)$, we have

$$(Q \otimes I_{N_0})^T \cdot \mathcal{L}(x,d) \cdot (Q \otimes I_{N_0}) \;=\; \mathcal{A}\Big(Q^T \cdot L_a(x,d) \cdot Q\Big) \;=\; \mathcal{A}(\tilde{L}). \tag{3.6}$$

Hence $\mathcal{L}(x,d) \succeq 0$ if and only if the $N_1 N_2 \cdots N_m$ blocks (3.5) are simultaneously positive semidefinite for all index sequences $j_1, j_2, \ldots, j_m$. Given that the $\lambda_{j_k}^{(k)}(x)$ are the eigenvalues of $B_k(x)$ we conclude that $\mathcal{L}(x,d) \succeq 0$ if and only if

$$d \cdot A_0 + \sum_{k=1}^{m} t_k A_k \succeq 0 \quad \text{for all } (t_1, \ldots, t_k) \text{ such that } t_k \text{ is eigenvalue of } B_k(x).$$

Since $d \cdot A_0 + \sum_{k=1}^{m} t_k A_k \succeq 0$ describes a convex set in t-space, we can now replace the condition "t_k is an eigenvalue of $B_k(x)$" by the equivalent condition

$$\lambda_{\min}(B_k(x)) \le t_k \le \lambda_{\max}(B_k(x)) \quad \text{for } k = 1, 2, \ldots, m.$$

We conclude that $\mathcal{L}(x,d) \succeq 0$ if and only if $(x,d) \in \mathcal{C}$. $\square$

The transformation from the given symmetric matrices A_i and $B_j(x)$ to the bigger matrix $\mathcal{L}(x,d)$ is easy to implement in `Matlab`, `Maple` or `Mathematica`. Here is a small explicit numerical example which illustrates this transformation.

Example 3.4. *Let $n = m = N_0 = N_1 = N_2 = 2$ and consider the input matrices*

$$A_0 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \qquad \textit{and}$$

$$B_1(x) = \begin{bmatrix} 3 - x_1 + 2x_2 & 2x_1 - x_2 - 2 \\ 2x_1 - x_2 - 2 & -1 + 2x_1 \end{bmatrix}, \; B_2(x) = \begin{bmatrix} 2 + x_1 & 1 + 3x_1 - x_2 \\ 1 + 3x_1 - x_2 & 3 - 2x_1 + x_2 \end{bmatrix}.$$

Then the three-dimensional spectrahedron $\mathcal{C}$ is represented by the LMI

$$\begin{bmatrix}
2d+3-x_1+2x_2 & d+5+2x_2 & 0 & 1+3x_1-x_2 & -2+2x_1-x_2 & -2+2x_1-x_2 & 0 & 0 \\
d+5+2x_2 & 2d+2+x_1 & 1+3x_1-x_2 & 1+3x_1-x_2 & -2+2x_1-x_2 & 0 & 0 & 0 \\
0 & 1+3x_1-x_2 & 2d+3-x_1+2x_2 & d+6-3x_1+3x_2 & 0 & 0 & -2+2x_1-x_2 & -2+2x_1-x_2 \\
1+3x_1-x_2 & 1+3x_1-x_2 & d+6-3x_1+3x_2 & 2d+3-2x_1+x_2 & 0 & 0 & -2+2x_1-x_2 & 0 \\
-2+2x_1-x_2 & -2+2x_1-x_2 & 0 & 0 & 2d-1+2x_1 & d+1+3x_1 & 0 & 1+3x_1-x_2 \\
-2+2x_1-x_2 & 0 & 0 & 0 & d+1+3x_1 & 2d+2+x_1 & 1+3x_1-x_2 & 1+3x_1-x_2 \\
0 & 0 & -2+2x_1-x_2 & -2+2x_1-x_2 & 0 & 1+3x_1-x_2 & 2d-1+2x_1 & d+2+x_2 \\
0 & 0 & -2+2x_1-x_2 & 0 & 1+3x_1-x_2 & 1+3x_1-x_2 & d+2+x_2 & 2d+3-2x_1+x_2
\end{bmatrix} \succeq 0.$$

Note that the boundary of $\mathcal{C}$ is a surface of degree eight in $\mathbb{R}^3$. If we fix a positive real value for d, then this LMI describes a two-dimensional spectrahedron. $\square$

4. Algebraic degree of the boundary. We have seen in Theorem 3.2 that the set $\mathcal{C}$ is a spectrahedron. This implies the property of *rigid convexity*, which was introduced by Helton and Vinnikov [8]. We briefly review this concept, starting with a discussion of *real zero* polynomials. Let $f(z)$ be a polynomial in s variables $(z_1, \ldots, z_s)$ and w a point in $\mathbb{R}^s$. We say $f(z)$ is a *real zero (RZ) polynomial with respect to* w if the univariate polynomial

$$g(\alpha) \;:=\; f(w + \alpha \cdot v)$$

has only real roots for any nonzero vector $v \in \mathbb{R}^n$. Being real zero with respect to w is a necessary condition [8] for $f(z)$ to have a determinantal representation

$$f(z) \;=\; \det(F_0 + z_1 F_1 + \cdots + z_s F_s),$$

where the F_i are constant symmetric matrices such that $F_0 + w_1 F_1 + \cdots + w_s F_s \succ 0$. Helton and Vinnikov [8] showed that the converse is true when $s = 2$, but a similar converse is not known for $s > 2$. For representations of convex sets as projections of spectrahedra in higher dimensional space we refer to [6, 7].

A convex set Γ is called *rigid convex* if Γ is a connected component of the set $\{z \in \mathbb{R}^s : f(z) > 0\}$ for some polynomial $f(z)$ that is real zero with respect to some interior point of Γ. As an example, the unit ball $\{z \in \mathbb{R}^n : \|z\| \leq 1\}$ is a rigid convex set. Note that, not every convex semialgebraic set is rigid convex. For instance, the set $\{z \in \mathbb{R}^2 : z_1^4 + z_2^4 \leq 1\}$ is not rigid convex as shown in [8].

The boundary of the 3-ellipse in Figure 2.1 is rigid convex of degree eight, and the curve (for fixed d) in Example 3.4 is also a rigid convex curve of degree eight. In light of Theorem 3.2, our discussion implies the following corollary.

Corollary 4.1. *The convex semialgebraic set $\mathcal{C} \subset \mathbb{R}^{n+1}$ in (1.1) is rigid convex.*

We now examine the algebraic properties of the boundary of the rigid convex set $\mathcal{C}$, starting from the explicit LMI representation $\mathcal{L}(x, d) \succeq 0$ which was constructed in (3.2). Since the tensor product is a bilinear operation, the entries of the symmetric matrix $\mathcal{L}(x, d)$ are linear in the $n+1$ variables (x, d), and they also depend linearly on the entries of the constant matrices A_i and $B_j^{(k)}$. We regard these entries as unknown parameters, and we let $K = \mathbb{Q}(A, B)$ denote the field extension they generate over the rational numbers $\mathbb{Q}$. Then $K[x, d]$ denotes the ring of polynomials in the $n + 1$ variables (x, d) with coefficients in the rational function field $K = \mathbb{Q}(A, B)$. Our object of interest is the determinant

$$r(x, d) \;=\; \det \mathcal{L}(x, d).$$

This is an element of $K[x, d]$. By a specialization of $r(x, d)$ we mean the image of $r(x, d)$ under any field homomorphism $K \to \mathbb{R}$. Thus a *specialization* of $r(x, d)$ is a polynomial in $\mathbb{R}[x, y]$ which arises as the determinant of the LMI representation (1.2) for some specific matrices A_i and $B_j^{(k)}$ with entries in $\mathbb{R}$.

Theorem 4.2. *The polynomial $r(x, d)$ has degree $N_0 N_1 \cdots N_m$, and it is irreducible as an element of $K[x, y]$. Generic specializations of $r(x, d)$ are irreducible polynomials of the same degree in $\mathbb{R}[x, d]$. Therefore, if the given symmetric matrices A_i and $B_j^{(k)}$ are generic, then the boundary of the spectrahedron $\mathcal{C}$ is a connected component of an irreducible real hypersurface of degree $N_0 N_1 \cdots N_m$.*

Proof. As in (3.4) we write $\lambda_1^{(k)}(x), \ldots, \lambda_{N_k}^{(k)}(x)$ for the eigenvalues of the matrix $B_k(x)$. These eigenvalues are elements in the algebraic closure of the rational function field $\mathbb{Q}(B,x)$. Here the entries of the matrix coefficient of x_i in $B_k(x)$ are regarded as variables over $\mathbb{Q}$. This implies that the characteristic polynomial of $B_k(x)$ is irreducible over $\mathbb{Q}(B,x)$ and the Galois group [10] of the corresponding algebraic field extension $\mathbb{Q}(B,x) \subset \mathbb{Q}(B,x)(\lambda_1^{(k)}(x), \ldots, \lambda_{N_k}^{(k)}(x))$ is the full symmetric group S_{N_k} on N_k letters. Let L be the algebraic field extension of $\mathbb{Q}(B,x)$ generated by all eigenvalues $\lambda_i^{(k)}(x)$ for $k = 1, \ldots, m$ and $i = 1, \ldots, N_k$. Since all scalar matrix entries are independent indeterminates over $\mathbb{Q}$, we conclude that $\mathbb{Q}(B,x) \subset L$ is an algebraic field extension of degree $N_1 \cdots N_m$, and the Galois group of this field extension is the product of symmetric groups $S_{N_1} \times \cdots \times S_{N_m}$.

The square matrix $\mathcal{L}(x,d)$ has $N_0 N_1 \cdots N_m$ rows and columns, and each entry is a linear polynomial in (x,d) with coefficients in $\mathbb{Q}(A,B)$. The polynomial $r(x,d)$ is the determinant of this matrix, so it has degree $N_0 N_1 \cdots N_m$ in (x,d). To see that it is irreducible, we argue as follows. In light of (3.6), the matrices $\mathcal{L}(x,d)$ and $\mathcal{A}(\tilde{L})$ have the same determinant, and this is the product of the determinants of the $N_1 N_2 \cdots N_m$ blocks (3.5) of size $N_0 \times N_0$. We conclude

$$r(x,d) \;=\; \prod_{j_1=1}^{N_1} \prod_{j_2=1}^{N_2} \cdots \prod_{j_m=1}^{N_m} \det\bigl(d \cdot A_0 \,+\, \lambda_{j_1}^{(1)}(x) A_1 + \lambda_{j_2}^{(2)}(x) A_2 + \cdots + \lambda_{j_m}^{(m)}(x) A_m\bigr).$$

The Galois group mentioned above acts transitively by permuting the factors in this product. No subset of the factors is left invariant under this permutation action. This proves that $r(x,d)$ is irreducible as univariate polynomial in $\mathbb{Q}(A,B,x)[d]$, and hence also as an $(m+1)$-variate polynomial in $\mathbb{Q}(A,B)[x,d]$.

The assertion in the third sentence follows because the property of a polynomial with parametric coefficients to be irreducible is Zariski open in the parameters. If we specialize the entries of A_i and $B_j^{(k)}$ to be random real numbers, then the resulting specialization of $r(x,d)$ is an irreducible polynomial in $\mathbb{R}[x,d]$. $\square$

Naturally, in many applications the entries of A_i and $B_j^{(k)}$ will not be generic but they have a special structure. In those cases, the characteristic polynomial of $B_k(x)$ may not be irreducible, and a more careful analysis is required in determining the degree of the hypersurface bounding the spectrahedron $\mathcal{C}$. Let $h_k(\alpha)$ be the univariate polynomial of minimal degree in $\mathbb{Q}(B,x)[\alpha]$ such that

$$h_k(\lambda_{\min}(B_k(x))) \;=\; h_k(\lambda_{\max}(B_k(x))) \;=\; 0.$$

holds for all points x in some open subset of $\mathbb{R}^n$. Thus, $h_k(\alpha)$ is the factor of the characteristic polynomial $\det(B_k(x) - \alpha I_{N_k})$ which is relevant for our problem. By modifying the argument in the previous proof, we obtain the following result.

Theorem 4.3. *The Zariski closure of the boundary of the spectrahedron $\mathcal{C}$ is a possibly reducible real hypersurface in $\mathbb{R}^{n+1}$. If the matrix A_0 is invertible then the degree of the polynomial in (x,d) which defines this hypersurface equals*

$$N_0 \cdot \prod_{k=1}^{m} \mathrm{degree}(h_k) \tag{4.1}$$

It is important to note that, for fixed d, the degree of the resulting polynomial in x can be smaller than (4.1). Thus this number is only an upper bound for the algebraic degree of the boundary of the spectrahedron $\{x \in \mathbb{R}^n : \mathcal{L}(x,d) \geq 0\}$.

Example 4.4. *The classical ellipse consists of all points $(x_1, x_2) \in \mathbb{R}^2$ whose sum of distances to two given points (u_i, v_i) is d. In the formulation sketched in Section 2 and worked out in detail in [12], we obtain $N_0 = 1$ and $N_i = \text{degree}(h_k) = 2$ for $k = 1, 2$. Hence the number (4.1) is four, and this is indeed the degree of the surface $\partial\mathcal{C} \subset \mathbb{R}^3$. Yet, the boundary of the ellipse is a curve of degree two.* $\square$

5. Applications and questions. The matrix cube defined in (1.1) has important applications in robust control [3, 4, 11]. Consider nonlinear feedback synthesis for the discrete dynamical system

$$\left.\begin{array}{rcll} x(k+1) & = & F(u(k)) \cdot x(k), & x(k) \in \mathbb{R}^n \\ u(k+1) & = & f(x(k))), & u(k) \in \mathbb{R}^m \end{array}\right\} \tag{5.1}$$

where $x(k)$ is the state and $u(k)$ is the feedback. We assume that the matrix $F(u) = F_0 + u_1 F_1 + \cdots + u_m F_m \in \mathbb{R}^{r\times s}$ is linear in u and that $f(x)$ is an algebraic function of $x = (x_1, \ldots, x_m)$. Such a function $f(x)$ can be implicitly defined by eigenvalues of linear matrices $B_k(x)$, that is, the coordinates of $u = f(x)$ satisfy

$$\det\big(u_k \cdot I_{N_k} - B_k(x)\big) = 0, \quad k = 1, \ldots, m.$$

In stability analysis one wish to know what states make the system (5.1) stable. More precisely, the task in stability analysis is to identify states x such that the feedback u satisfies $\|F(u)\|_2 \le 1$. Since this inequality is equivalent to the LMI

$$\begin{bmatrix} I_r & F(u) \\ F(u)^T & I_s \end{bmatrix} \preceq I_{r+s}, \tag{5.2}$$

the task can be reformulated as follows: find state vectors x satisfying (5.2)

$$\text{for all } u \in \prod_{k=1}^{m} \left[\lambda_{min}(B_k(x)), \lambda_{max}(B_k(x))\right].$$

Therefore we conclude that such states x are the elements of a convex set of the form (1.1) when every $B_k(x)$ can be chosen to be symmetric. In particular, Theorem 3.2 furnishes an LMI representation for the set of stable states.

In Section 2 we have seen that the LMI representation of multifocal ellipses in [12] arise as a special case of our construction. We next propose a natural matrix-theoretic generalization of this, namely, we shall define the *matrix ellipsoid* and the *matrix m-ellipsoid*. Recall that an m-ellipsoid is the subset of $\mathbb{R}^n$ defined as

$$\Big\{x \in \mathbb{R}^n : a_1\|x - u_1\|_2 + \cdots + a_m\|x - u_m\|_2 \le d\Big\}$$

for some constants $a_1, a_2, \ldots, a_m, d > 0$ and fixed foci $u_1, u_2, \cdots, u_m \in \mathbb{R}^n$. We define the *matrix m-ellipsoid* to be the convex set

$$\Big\{x \in \mathbb{R}^n : A_1\|x - u_1\|_2 + \cdots + A_m\|x - u_m\|_2 \preceq d \cdot I_N\Big\}, \tag{5.3}$$

for some $d > 0$, fixed foci $u_1, u_2, \cdots, u_m \in \mathbb{R}^n$, and positive definite symmetric N-by-N matrices $A_1, \ldots, A_m \succ 0$. To express the matrix m-ellipsoid (5.3) as an instance of the spectrahedron $\mathcal{C}$ in (1.1), we fix $A_0 = I_N$ and the linear matrices

$$B_k(x) \quad = \quad \begin{bmatrix} 0 & x_1 - u_{k,1} & \cdots & x_n - u_{k,n} \\ x_1 - u_{k,1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_n - u_{k,n} & 0 & \cdots & 0 \end{bmatrix}.$$

Theorem 3.2 provides an LMI representation of the matrix m-ellipsoid (5.3). This generalizes the LMI given in [12, §4.2] for the case when all A_i are scalars. By Theorem 4.3, the boundary of (5.3) is a hypersurface of degree at most $N2^m$.

One obvious generalization of our parametric matrix cube problem is the set

$$\tilde{C} := \left\{ (x,d) \;\middle|\; \begin{array}{r} d \cdot A_0 + \sum_{k=1}^{m} t_k A_k \succeq 0, \, \forall\, (t_1, \cdots, t_m) : \\ \lambda_{\min}(B_k(x)) \le t_k \le \lambda_{\max}(E_k(x)) \\ \text{for } k = 1, \cdots, m \end{array} \right\}$$

where $E_k(x)$ are linear symmetric matrices different from $B_k(x)$ for $k = 1, \ldots, m$. Since the minimum eigenvalue function $\lambda_{\min}(\,\cdot\,)$ is concave and the maximum eigenvalue function $\lambda_{\max}(\,\cdot\,)$ is convex, we see that the set $\tilde{C}$ defined as above is also a convex set. Assuming the extra hypotheses $\lambda_{\max}(E_k(x)) \ge \lambda_{\max}(B_k(x))$ and $\lambda_{\min}(E_k(x)) \ge \lambda_{\min}(B_k(x))$, the convex set $\mathcal{C}$ can be equivalently defined as

$$\left\{ (x,d) \;\middle|\; \begin{array}{r} d \cdot A_0 + \sum_{k=1}^{m} t_k A_k \succeq 0, \, \forall\, (t_1, \cdots, t_m) : \\ \lambda_{\min}(D_k(x)) \le t_k \le \lambda_{\max}(D_k(x)) \\ \text{for } k = 1, \cdots, m \end{array} \right\},$$

where $D_k(x) = \operatorname{diag}\big(B_k(x), E_k(x)\big)$ is a block diagonal matrix. Therefore, $\tilde{C}$ is a special case of (1.1) and Theorem 3.2 furnishes an LMI representation. We do not know whether the extra hypotheses are necessary for $\tilde{C}$ to be a spectrahedron.

An interesting research problem concerning the matrix cube (1.1) is to find the smallest size of an LMI representation. This question was also raised in [12, §5] for the case of m-ellipses. Theorem 3.2 gives an LMI representation of size $N_0 N_1 \cdots N_m$. The degree of the boundary $\partial\mathcal{C}$ is given by the formula $N_0 \prod_{k=1}^{m} \text{degree}(h_k)$, by Theorem 4.3. If degree(h_k) is smaller than N_k, then the size of the LMI in Theorem 3.2 exceeds the degree of $\partial\mathcal{C}$. In this situation, is it possible to find an LMI for $\mathcal{C}$ with size smaller than $N_0 N_1 \cdots N_m$? When $n = 2$ and d is fixed, the projection of $\mathcal{C}$ into x-space is a two dimensional convex set $\mathcal{C}_x$ described by one LMI of size $N_0 N_1 \cdots N_m$. The work of Helton and Vinnikov [8] shows that there exists an LMI for $\mathcal{C}_x$ having size equal to the degree of $\mathcal{C}_x$. How can the LMI of Theorem 3.2 be transformed into such a minimal LMI?

REFERENCES

[1] S. Basu, R. Pollack and M.-F. Roy. *Algorithms in Real Algebraic Geometry*. Springer-Verlag, Berlin, 2006.

[2] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM Studies in Applied Mathematics, Vol. **15**, Society for Industrial and Applied Math., Philadelphia, 1994

[3] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research* **23** (1998), No.4, 769–805.

[4] A. Ben-Tal, A. Nemirovski and C. Roos. Extended matrix cube theorems with applications to μ-theory in control. *Mathematics of Operations Research* **28** (2003), 497-523.

[5] B.-D. Chen and S. Lall. Degree bounds for polynomial verification of the matrix cube problem. *Proceedings of the IEEE Conference on Decision and Control*, pp. 4405–4410, 2006.

[6] J.W. Helton and J. Nie. Semidefinite representation of convex sets. Preprint, 2007. `arXiv:0705.4068`.

[7] J.W. Helton and J. Nie. Sufficient and necessary conditions for semidefinite representability of convex hulls and sets. Preprint, 2007. `arxiv:0709.4017`.

[8] J.W. Helton and V. Vinnikov. Linear matrix inequality representation of sets. *Comm. Pure Appl. Math.* **60** (2007), no. 5, 654–674.
[9] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1994.
[10] S. Lang. *Algebra*. Third edition, Springer-Verlag, New York, 2000.
[11] A. Nemirovski. *The Matrix Cube Problem: Approximations and Applications*. Lecture Notes, 2001, `http://www2.isye.gatech.edu/~nemirovs/`.
[12] J. Nie, P. Parrilo and B. Sturmfels. Semidefinite Representation of the k-Ellipse. *IMA Volume 146: Algorithms in Algebraic Geometry (Eds. A. Dickenstein, F.-O. Schreyer, and A. Sommese)*, Springer, New York, 2008.
[13] J. Nie, K. Ranestad and B. Sturmfels, The algebraic degree of semidefinite programming. `arXiv:math/0611562`, to appear in *Math. Programming*.
[14] M. Ramana and A.J. Goldman. Some geometric results in semidefinite programming. *Journal of Global Optimization* **7** (1995) 33–50.
[15] J.F. Sturm. SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, **11/12** (1999) 625–653.
[16] H. Wolkowicz, R. Saigal, and L. Vandenberghe (Eds.). *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, Kluwer Academic Publishers, 2000.