

NEW APPROXIMATIONS FOR THE CONE OF COPOSITIVE MATRICES AND ITS DUAL

JEAN B. LASSERRE

ABSTRACT. We provide convergent hierarchies for the convex cone \mathcal{C} of copositive matrices and its dual \mathcal{C}^* , the cone of completely positive matrices. In both cases the corresponding hierarchy consists of nested spectrahedra and provide outer (resp. inner) approximations for \mathcal{C} (resp. for its dual \mathcal{C}^*), thus complementing previous inner (resp. outer) approximations for \mathcal{C} (for \mathcal{C}^*). In particular, both inner and outer approximations have a very simple interpretation. Finally, extension to \mathcal{K} -copositivity and \mathcal{K} -complete positivity for a closed convex cone \mathcal{K} , is straightforward.

1. INTRODUCTION

In recent years the convex cone \mathcal{C} of *copositive* matrices and its dual cone \mathcal{C}^* of *completely positive* matrices have attracted a lot of attention, in part because several interesting NP-hard problems can be modelled as convex conic optimization problems over those cones. For a survey of results and discussion on \mathcal{C} and its dual, the interested reader is referred to e.g. Anstreicher and Burer [1], Bomze [2], Bomze et al. [3], Burer [5], Dür [8] and Hiriart-Urruty and Seeger [11].

As optimizing over \mathcal{C} (or its dual) is in general difficult, a typical approach is to optimize over simpler and more tractable cones. In particular, nested hierarchies of tractable convex cones \mathcal{C}_k , $k \in \mathbb{N}$, that provide *inner* approximations of \mathcal{C} have been proposed, notably by Parrilo [13], deKlerk and Pasechnik [7], Bomze and deKlerk [4], as well as Peña et al. [14]. For example, denoting by \mathcal{N} (resp. \mathcal{S}_+) the convex cone of nonnegative (resp. positive semidefinite) matrices, the first cone in the hierarchy of [7] is \mathcal{N} , and $\mathcal{N} + \mathcal{S}_+$ in that of [13], whereas the hierarchy of [14] is in sandwich between that of [7] and [13]. Of course, to each such hierarchy of inner approximations (\mathcal{C}_k) , $k \in \mathbb{N}$, of \mathcal{C} , one may associate the hierarchy (\mathcal{C}_k^*) , $k \in \mathbb{N}$, of dual cones which provides outer approximations of \mathcal{C}^* . However, quoting Dür in [8]: “We are not aware of comparable approximation schemes that approximate the completely positive cone (i.e. \mathcal{C}^*) from the interior.”

Let us also mention Gaddum’s characterization of copositive matrices described in Gaddum [9] for which the set membership problem “ $\mathbf{A} \in \mathcal{C}$ ” (with \mathbf{A} known) reduces to solving a linear programming (LP) problem (whose size is not polynomially bounded in the input size of the matrix \mathbf{A}). However, Gaddum’s characterization is *not* convex in the coefficients of the matrix \mathbf{A} , and so cannot be used to provide a hierarchy of outer approximations of \mathcal{C} . On the other hand, de Klerk and Pasechnik [6] have used Gaddum’s characterization to help solve quadratic optimization

1991 *Mathematics Subject Classification.* 15B48 90C22.

Key words and phrases. copositive matrices; completely positive matrices; semidefinite relaxations.

problems on the simplex via solving a hierarchy of LP-relaxations of increasing size and with *finite* convergence of the process.

Contribution. The contribution of this note is precisely to describe an explicit hierarchy of tractable convex cones that provide *outer* approximations of \mathcal{C} which converge monotonically and asymptotically to \mathcal{C} . And so, by duality, the corresponding hierarchy of dual cones provides *inner* approximations of \mathcal{C}^* converging to \mathcal{C}^* , answering Dür's question and also showing that \mathcal{C} and \mathcal{C}^* can be sandwiched between two converging hierarchies of tractable convex cones. The present outer approximations are *not* polyhedral and hence are different from the outer polyhedral approximations \mathcal{O}_r^n of \mathcal{C} defined in Yildirim [20]. \mathcal{O}_r^n are based on a certain discretization $\Delta(n, r)$ of the simplex $\Delta_n := \{\mathbf{x} \in \mathbb{R}_+^n : \mathbf{e}^T \mathbf{x} = 1\}$ whose size is parametrized by r . All matrices associated with quadratic forms nonnegative on $\Delta(n, r)$ belong to \mathcal{O}_n^r for all r , and $\mathcal{C} = \bigcap_{r=1}^{\infty} \mathcal{O}_n^r$; Combining with the inner approximations of de Klerk and Pasechnik [7], the author provides tight bounds on the gap between upper and lower bounds for quadratic optimization problems; for more details the interested reader is referred to Yildirim [20].

In fact, our result is a specialization of a more general result of [12] about nonnegativity of polynomials on a closed set $\mathbf{K} \subset \mathbb{R}^n$, in the specific context of quadratic forms and $\mathbf{K} = \mathbb{R}_+^n$. In such a context, and identifying copositive matrices with quadratic forms nonnegative on $\mathbf{K} = \mathbb{R}_+^n$, this specialization yields inner approximations for \mathcal{C}^* with a very simple interpretation directly related to the definition of \mathcal{C}^* , which might be of interest for the community interested in \mathcal{C} and \mathcal{C}^* but might be “lost” in the general case treated in [12], whence the present note. Finally, following Burer [5], and $\mathcal{K} \subset \mathbb{R}^n$ being a closed convex cone, one may also consider the convex cone $\mathcal{C}_{\mathcal{K}}$ of \mathcal{K} -copositive matrices, i.e., real symmetric matrices \mathbf{A} such that $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ on \mathcal{K} , and its dual cone $\mathcal{C}_{\mathcal{K}}^*$ of \mathcal{K} -completely positive matrices. Then the outer approximations previously defined have an immediate and straightforward analogue (as well as the inner approximations of the dual).

2. MAIN RESULT

2.1. Notation and definitions. Let $\mathbb{R}[\mathbf{x}]$ be the ring of polynomials in the variables $\mathbf{x} = (x_1, \dots, x_n)$. Denote by $\mathbb{R}[\mathbf{x}]_d \subset \mathbb{R}[\mathbf{x}]$ the vector space of polynomials of degree at most d , which forms a vector space of dimension $s(d) = \binom{n+d}{d}$, with e.g., the usual canonical basis (\mathbf{x}^α) of monomials. Also, let $\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n : \sum_i \alpha_i \leq d\}$ and denote by $\Sigma[\mathbf{x}] \subset \mathbb{R}[\mathbf{x}]$ (resp. $\Sigma[\mathbf{x}]_d \subset \mathbb{R}[\mathbf{x}]_{2d}$) the space of sums of squares (s.o.s.) polynomials (resp. s.o.s. polynomials of degree at most $2d$). If $f \in \mathbb{R}[\mathbf{x}]_d$, write $f(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha \mathbf{x}^\alpha$ in the canonical basis and denote by $\mathbf{f} = (f_\alpha) \in \mathbb{R}^{s(d)}$ its vector of coefficients. Finally, let \mathcal{S}^n denote the space of $n \times n$ real symmetric matrices, with inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \text{trace } \mathbf{A} \mathbf{B}$, and where the notation $\mathbf{A} \succeq 0$ (resp. $\mathbf{A} \succ 0$) stands for \mathbf{A} is positive semidefinite.

Given $\mathbf{K} \subseteq \mathbb{R}^n$, denote by $\text{cl } \mathbf{K}$ (resp. $\text{conv } \mathbf{K}$) the closure (resp. the convex hull) of \mathbf{K} . Recall that given a convex cone $\mathbf{K} \subseteq \mathbb{R}^n$, the convex cone $\mathbf{K}^* = \{\mathbf{y} \in \mathbb{R}^n : \langle \mathbf{y}, \mathbf{x} \rangle \geq 0 \forall \mathbf{x} \in \mathbf{K}\}$ is called the *dual cone* of \mathbf{K} , and satisfies $(\mathbf{K}^*)^* = \text{cl } \mathbf{K}$. Moreover, given two convex cones $\mathbf{K}_1, \mathbf{K}_2 \subseteq \mathbb{R}^n$,

$$\begin{aligned} \mathbf{K}_1^* \cap \mathbf{K}_2^* &= (\mathbf{K}_1 + \mathbf{K}_2)^* = (\mathbf{K}_1 \cup \mathbf{K}_2)^* \\ (\mathbf{K}_1 \cap \mathbf{K}_2)^* &= \text{cl } (\mathbf{K}_1^* + \mathbf{K}_2^*) = \text{cl } (\text{conv } (\mathbf{K}_1^* \cup \mathbf{K}_2^*)). \end{aligned}$$

See for instance Rockafellar [15, Theorem 3.8; Corollary 16.4.2].

Moment matrix. With a sequence $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, let $L_{\mathbf{y}} : \mathbb{R}[\mathbf{x}] \rightarrow \mathbb{R}$ be the linear functional

$$h \quad (= \sum_{\alpha} h_{\alpha} \mathbf{x}^{\alpha}) \quad \mapsto \quad L_{\mathbf{y}}(h) = \sum_{\alpha} h_{\alpha} y_{\alpha}, \quad h \in \mathbb{R}[\mathbf{x}].$$

With $d \in \mathbb{N}$, let $\mathbf{M}_d(\mathbf{y})$ be the symmetric matrix with rows and columns indexed in \mathbb{N}_d^n , and defined by:

$$(2.1) \quad \mathbf{M}_d(\mathbf{y})(\alpha, \beta) := L_{\mathbf{y}}(\mathbf{x}^{\alpha+\beta}) = y_{\alpha+\beta}, \quad (\alpha, \beta) \in \mathbb{N}_d^n \times \mathbb{N}_d^n.$$

The matrix $\mathbf{M}_d(\mathbf{y})$ is called the moment matrix associated with \mathbf{y} , and it is straightforward to check that

$$[L_{\mathbf{y}}(g^2) \geq 0 \quad \forall g \in \mathbb{R}[\mathbf{x}]] \quad \Leftrightarrow \quad \mathbf{M}_d(\mathbf{y}) \succeq 0, \quad d = 0, 1, \dots$$

Localizing matrix. Similarly, with $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, and $f \in \mathbb{R}[\mathbf{x}]$ written

$$\mathbf{x} \mapsto f(\mathbf{x}) = \sum_{\gamma \in \mathbb{N}^n} f_{\gamma} \mathbf{x}^{\gamma},$$

let $\mathbf{M}_d(f \mathbf{y})$ be the symmetric matrix with rows and columns indexed in \mathbb{N}_d^n , and defined by:

$$(2.2) \quad \mathbf{M}_d(f \mathbf{y})(\alpha, \beta) := L_{\mathbf{y}}(f(\mathbf{x}) \mathbf{x}^{\alpha+\beta}) = \sum_{\gamma} f_{\gamma} y_{\alpha+\beta+\gamma}, \quad (\alpha, \beta) \in \mathbb{N}_d^n \times \mathbb{N}_d^n.$$

The matrix $\mathbf{M}_d(f \mathbf{y})$ is called the localizing matrix associated with \mathbf{y} and $f \in \mathbb{R}[\mathbf{x}]$. Observe that

$$(2.3) \quad \langle \mathbf{g}, \mathbf{M}_d(f \mathbf{y}) \mathbf{g} \rangle = L_{\mathbf{y}}(g^2 f), \quad \forall \mathbf{g} \in \mathbb{R}[\mathbf{x}]_d,$$

and so if \mathbf{y} has a representing finite Borel measure μ , i.e., if

$$y_{\alpha} = \int_{\mathbb{R}^n} \mathbf{x}^{\alpha} d\mu, \quad \forall \alpha \in \mathbb{N}^n,$$

then (2.3) reads

$$(2.4) \quad \langle \mathbf{g}, \mathbf{M}_d(f \mathbf{y}) \mathbf{g} \rangle = L_{\mathbf{y}}(g^2 f) = \int_{\mathbb{R}^n} g(\mathbf{x})^2 f(\mathbf{x}) d\mu(\mathbf{x}), \quad \forall \mathbf{g} \in \mathbb{R}[\mathbf{x}]_d.$$

Actually, the localizing matrix $\mathbf{M}_d(f \mathbf{y})$ is nothing less than the moment matrix associated with the sequence $\mathbf{z} = f \mathbf{y} = (z_{\alpha})$, $\alpha \in \mathbb{N}^n$, with $z_{\alpha} = \sum_{\gamma} f_{\gamma} y_{\alpha+\gamma}$. In particular, if f is nonnegative on the support $\text{supp } \mu$ of μ then the localizing matrix $\mathbf{M}_d(f \mathbf{y})$ is just the moment matrix associated with the finite Borel measure $d\mu_f = f d\mu$, absolutely continuous with respect to μ (denoted $\mu_f \ll \mu$), and with density f . For instance, with $d = 1$ one may write

$$\mathbf{M}_1(f \mathbf{y}) = \int_{\mathbb{R}^n} \begin{bmatrix} 1 & | & \mathbf{x}^T \\ - & & - \\ \mathbf{x} & | & \mathbf{x}\mathbf{x}^T \end{bmatrix} f(\mathbf{x}) d\mu(\mathbf{x}) = \int_{\mathbb{R}^n} \begin{bmatrix} 1 & | & \mathbf{x}^T \\ - & & - \\ \mathbf{x} & | & \mathbf{x}\mathbf{x}^T \end{bmatrix} d\mu_f(\mathbf{x}),$$

or, equivalently,

$$(2.5) \quad \mathbf{M}_1(f \mathbf{y}) = \text{mass}(\mu_f) \times \begin{bmatrix} 1 & | & \mathbf{E}_{\tilde{\mu}_f}(\mathbf{x})^T \\ - & & - \\ \mathbf{E}_{\tilde{\mu}_f}(\mathbf{x}) & | & \mathbf{E}_{\tilde{\mu}_f}(\mathbf{x}\mathbf{x}^T) \end{bmatrix},$$

where $\mathbf{E}_{\tilde{\mu}_f}(\cdot)$ denotes the expectation operator associated with the normalization $\tilde{\mu}_f$ of μ_f (as a probability measure), and $\mathbf{E}_{\tilde{\mu}_f}(\mathbf{x}\mathbf{x}^T)$ denotes the matrix of noncentral second-order moments of $\tilde{\mu}_f$ (and the covariance matrix of $\tilde{\mu}_f$ if $\mathbf{E}_{\tilde{\mu}_f}(\mathbf{x}) = 0$).

2.2. Main result. In [12, Theorem 3.2] the author has shown in a general context that a polynomial $f \in \mathbb{R}[\mathbf{x}]$ is nonnegative on a closed set $\mathbf{K} \subseteq \mathbb{R}^n$ if and only if

$$(2.6) \quad \int_{\mathbf{K}} g^2 f d\mu \geq 0, \quad \forall g \in \mathbb{R}[\mathbf{x}],$$

where μ is a given finite Borel measure with support $\text{supp } \mu$ being exactly \mathbf{K} ; if \mathbf{K} is compact then μ is arbitrary whereas if \mathbf{K} is not compact then μ has to satisfy a certain growth condition on its moments. If $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, is the sequence of moments of μ then (2.6) is in turn equivalent to

$$\mathbf{M}_d(f \mathbf{y}) \succeq 0, \quad \forall d = 0, 1, \dots,$$

where $\mathbf{M}_d(f \mathbf{y})$ is the localizing matrix associated with f and \mathbf{y} , defined in (2.2). In this section we particularize this result to the case of copositive matrices viewed as homogeneous forms of degree 2, nonnegative on the closed set $\mathbf{K} = \mathbb{R}_+^n$.

So with $\mathbf{A} = (a_{ij}) \in \mathcal{S}^n$, let denote by $f_{\mathbf{A}} \in \mathbb{R}[\mathbf{x}]_2$ the quadratic form $\mathbf{x} \mapsto \mathbf{x}^T \mathbf{A} \mathbf{x}$, and let μ be the joint probability measure associated with n *i.i.d.* exponential variates (with mean 1), with support $\text{supp } \mu = \mathbb{R}_+^n$, and with moments $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, given by:

$$(2.7) \quad y_\alpha = \int_{\mathbb{R}_+^n} \mathbf{x}^\alpha d\mu(\mathbf{x}) = \int_{\mathbb{R}_+^n} \mathbf{x}^\alpha \exp\left(-\sum_{i=1}^n x_i\right) d\mathbf{x} = \prod_{i=1}^n \alpha_i!, \quad \forall \alpha \in \mathbb{N}^n.$$

Recall that a matrix $\mathbf{A} \in \mathcal{S}^n$ is copositive if $f_{\mathbf{A}}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}_+^n$, and denote by $\mathcal{C} \subset \mathcal{S}^n$ the cone of copositive matrices, i.e.,

$$(2.8) \quad \mathcal{C} := \{ \mathbf{A} \in \mathcal{S}^n : f_{\mathbf{A}}(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}_+^n \}.$$

Its dual cone is the closed convex cone of *completely positive*, i.e., matrices of \mathcal{S}^n that can be written as the sum of finitely many rank-one matrices $\mathbf{x}\mathbf{x}^T$, with $\mathbf{x} \in \mathbb{R}_+^n$, i.e.,

$$(2.9) \quad \mathcal{C}^* = \text{conv} \{ \mathbf{x}\mathbf{x}^T : \mathbf{x} \in \mathbb{R}_+^n \}.$$

Next, introduce the following sets $\mathcal{C}_d \subset \mathcal{S}^n$, $d = 0, 1, \dots$, defined by:

$$(2.10) \quad \mathcal{C}_d := \{ \mathbf{A} \in \mathcal{S}^n : \mathbf{M}_d(f_{\mathbf{A}} \mathbf{y}) \succeq 0 \}, \quad d = 0, 1, \dots$$

where $\mathbf{M}_d(f_{\mathbf{A}} \mathbf{y})$ is the localizing matrix defined in (2.2), associated with the quadratic form $f_{\mathbf{A}}$ and the sequence \mathbf{y} in (2.7).

Observe that in view of the definition (2.2) of the localizing matrix, the entries of the matrix $\mathbf{M}_d(f_{\mathbf{A}} \mathbf{y})$ are homogeneous and linear in \mathbf{A} . Therefore, the condition $\mathbf{M}_d(f_{\mathbf{A}} \mathbf{y}) \succeq 0$ is a homogeneous Linear Matrix Inequality (LMI) and defines a closed convex cone (in fact a spectrahedron) of \mathcal{S}^n (or equivalently, of $\mathbb{R}^{n(n+1)/2}$). Each $\mathcal{C}_d \subset \mathcal{S}^n$ is a convex cone defined solely in terms of the entries (a_{ij}) of $\mathbf{A} \in \mathcal{S}^n$, and the hierarchy of spectrahedra (\mathcal{C}_d) , $d \in \mathbb{N}$, provides a nested sequence of outer approximations of \mathcal{C} .

Theorem 2.1. *Let \mathbf{y} be as in (2.7) and let $\mathcal{C}_d \subset \mathcal{S}^n$, $d = 0, 1, \dots$, be the hierarchy of convex cones defined in (2.10). Then $\mathcal{C}_0 \supset \mathcal{C}_1 \cdots \supset \mathcal{C}_d \cdots \supset \mathcal{C}$ and $\mathcal{C} = \bigcap_{d=0}^{\infty} \mathcal{C}_d$.*

The proof is a direct consequence of [12, Theorem 3.3] with $\mathbf{K} = \mathbb{R}_+^n$ and $f = f_{\mathbf{A}}$. Since $f_{\mathbf{A}}$ is homogeneous, alternatively one may use the probability measure ν uniformly supported on the n -dimensional simplex $\Delta = \{\mathbf{x} \in \mathbb{R}_+^n : \sum_i x_i \leq 1\}$ and invoke [12, Theorem 3.2]. The moments $\tilde{\mathbf{y}} = (\tilde{y}_\alpha)$ of ν are also quite simple to obtain and read:

$$(2.11) \quad \tilde{y}_\alpha = \int_{\Delta} \mathbf{x}^\alpha d\mathbf{x} = \frac{\alpha_1! \cdots \alpha_n!}{(n + \sum_i \alpha_i)!}, \quad \forall \alpha \in \mathbb{N}^n.$$

(See e.g. Grundmann [10].)

Observe that the set membership problem “ $\mathbf{A} \in \mathcal{C}_d$ ”, i.e., testing whether a given matrix $\mathbf{A} \in \mathcal{S}^n$ belongs to \mathcal{C}_d , is an eigenvalue problem as one has to check whether the smallest eigenvalue of $\mathbf{M}_d(f_{\mathbf{A}} \mathbf{y})$ is nonnegative. Therefore, instead of using standard packages for Linear Matrix Inequalities, one may use powerful specialized softwares for computing eigenvalues of real symmetric matrices.

We next describe an inner approximation of the convex cone \mathcal{C}^* via the hierarchy of convex cones (\mathcal{C}_d^*) , $d \in \mathbb{N}$, where each \mathcal{C}_d^* is the dual cone of \mathcal{C}_d in Theorem 2.1.

Recall that $\Sigma[\mathbf{x}]_d$ is the space of polynomials that are sums of squares of polynomials of degree at most d . A matrix $\mathbf{A} \in \mathcal{S}^n$ is also identified with a vector $\mathbf{a} \in \mathbb{R}^{n(n+1)/2}$, and conversely, with any vector $\mathbf{a} \in \mathbb{R}^{n(n+1)/2}$ is associated a matrix $\mathbf{A} \in \mathcal{S}^n$. For instance, with $n = 2$,

$$(2.12) \quad \mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \leftrightarrow \mathbf{a} = \begin{bmatrix} a \\ 2b \\ c \end{bmatrix}.$$

So we will not distinguish between a convex cone in $\mathbb{R}^{n(n+1)/2}$ and the corresponding cone in \mathcal{S}^n .

Theorem 2.2. *Let $\mathcal{C}_d \subset \mathcal{S}^n$ be the convex cone defined in (2.10). Then*

$$(2.13) \quad \mathcal{C}_d^* = \text{cl} \left\{ \left(\langle \mathbf{X}, \mathbf{M}_d(x_i x_j \mathbf{y}) \rangle \right)_{1 \leq i \leq j \leq n} : \mathbf{X} \in \mathcal{S}_+^{s(d)} \right\}.$$

Equivalently:

$$(2.14) \quad \begin{aligned} \mathcal{C}_d^* &= \text{cl} \left\{ \int_{\mathbb{R}_+^n} \mathbf{xx}^T \underbrace{\sigma(\mathbf{x})}_{d\mu_\sigma(\mathbf{x})} d\mu(\mathbf{x}) : \sigma \in \Sigma[\mathbf{x}]_d \right\} \\ &= \text{cl} \left\{ \text{mass}(\mu_\sigma) \mathbf{E}_{\tilde{\mu}_\sigma}(\mathbf{xx}^T) : \sigma \in \Sigma[\mathbf{x}]_d \right\}, \end{aligned}$$

with $\tilde{\mu}_\sigma$ and $\mathbf{E}_{\tilde{\mu}_\sigma}(\mathbf{xx}^T)$ as in (2.5).

Proof. Let

$$\Delta_d := \left\{ \left(\langle \mathbf{X}, \mathbf{M}_d(x_i x_j \mathbf{y}) \rangle \right)_{1 \leq i \leq j \leq n} : \mathbf{X} \in \mathcal{S}_+^{s(d)} \right\},$$

so that

$$\begin{aligned}
\Delta_d^* &= \left\{ \mathbf{a} \in \mathbb{R}^{n(n+1)/2} : \sum_{1 \leq i \leq j \leq n} a_{ij} \langle \mathbf{X}, \mathbf{M}_d(x_i x_j \mathbf{y}) \rangle \geq 0 \quad \forall \mathbf{X} \in \mathcal{S}_+^{s(d)} \right\}, \\
&= \left\{ \mathbf{a} \in \mathbb{R}^{n(n+1)/2} : \left\langle \mathbf{X}, \mathbf{M}_d \left(\left(\sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j \right) \mathbf{y} \right) \right\rangle \geq 0 \quad \forall \mathbf{X} \in \mathcal{S}_+^{s(d)} \right\}, \\
&= \{ \mathbf{A} \in \mathcal{S}^n : \langle \mathbf{X}, \mathbf{M}_d(f_{\mathbf{A}} \mathbf{y}) \rangle \geq 0 \quad \forall \mathbf{X} \in \mathcal{S}_+^{s(d)} \} \quad [\text{with } \mathbf{A}, \mathbf{a} \text{ as in (2.12)}] \\
&= \{ \mathbf{A} \in \mathcal{S}^n : \mathbf{M}_d(f_{\mathbf{A}} \mathbf{y}) \succeq 0 \} = \mathcal{C}_d.
\end{aligned}$$

And so we obtain the desired result $\mathcal{C}_d^* = (\Delta_d^*)^* = \text{cl}(\Delta_d)$. Next, writing the singular decomposition of \mathbf{X} as $\sum_{k=0}^s \mathbf{q}_k \mathbf{q}_k^T$ for some $s \in \mathbb{N}$ and some vectors $(\mathbf{q}_k) \subset \mathbb{R}^{s(d)}$, one obtains that for every $1 \leq i \leq j \leq n$,

$$\begin{aligned}
\langle \mathbf{X}, \mathbf{M}_d(x_i x_j \mathbf{y}) \rangle &= \sum_{k=0}^s \langle \mathbf{q}_k \mathbf{q}_k^T, \mathbf{M}_d(x_i x_j \mathbf{y}) \rangle = \sum_{k=0}^s \langle \mathbf{q}_k, \mathbf{M}_d(x_i x_j \mathbf{y}) \mathbf{q}_k \rangle \\
&= \sum_{k=0}^s \int_{\mathbb{R}_+^n} x_i x_j q_k(\mathbf{x})^2 d\mu(\mathbf{x}) \quad [\text{by (2.4)}] \\
&= \int_{\mathbb{R}_+^n} x_i x_j \underbrace{\sigma(\mathbf{x})}_{d\mu_\sigma(\mathbf{x})} d\mu(\mathbf{x}),
\end{aligned}$$

where $\sigma(\mathbf{x}) = \sum_{k=0}^s q_k(\mathbf{x})^2 \in \Sigma[\mathbf{x}]_d$, and $\mu_\sigma(B) = \int_B \sigma d\mu$ for all Borel sets B . \square

So Theorem 2.2 states that \mathcal{C}_d^* is the closure of the convex cone generated by second-order moments of measures $d\mu_\sigma = \sigma d\mu$, absolutely continuous with respect to μ (hence with support on \mathbb{R}_+^n) and with density being a s.o.s. polynomial σ of degree at most $2d$. Of course we immediately have:

Corollary 2.3. *Let \mathcal{C}_d^* , $d \in \mathbb{N}$, be as in (2.14). Then $\mathcal{C}_d^* \subset \mathcal{C}_{d+1}^*$ for all $d \in \mathbb{N}$, and*

$$\mathcal{C}^* = \text{cl} \bigcup_{d=0}^{\infty} \mathcal{C}_d^*.$$

Proof. As $\mathcal{C}_d^* \subset \mathcal{C}_{d+1}^*$ for all $d \in \mathbb{N}$, the result follows from

$$\mathcal{C}^* = \left(\bigcap_{d=0}^{\infty} \mathcal{C}_d \right)^* = \text{cl} \left(\text{conv} \bigcup_{d=0}^{\infty} \mathcal{C}_d^* \right) = \text{cl} \bigcup_{d=0}^{\infty} \mathcal{C}_d^*.$$

\square

In other words, \mathcal{C}_d^* approximates $\mathcal{C}^* = \text{conv} \{ \mathbf{x} \mathbf{x}^T : \mathbf{x} \in \mathbb{R}_+^n \}$ (i.e., the convex hull of second-order moments of Dirac measures with support in \mathbb{R}_+^n) from inside by second-order moments of measures $\mu_\sigma \ll \mu$ whose density is a s.o.s. polynomial σ of degree at most $2d$, and better and better approximations are obtained by letting d increase.

Example 1. For instance, with $n = 2$, and $\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, it is known that \mathbf{A} is copositive if and only if $a, c \geq 0$ and $b + \sqrt{ac} \geq 0$; see e.g. [11]. Let μ be

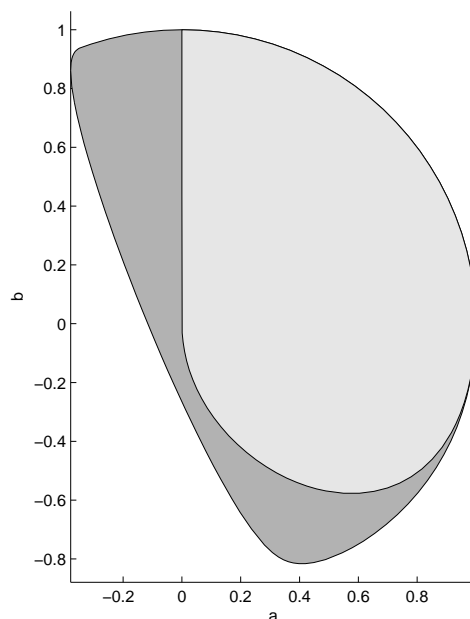


FIGURE 1. $n = 2$: Projection on the (a, b) -plane of \mathcal{C} versus \mathcal{C}_1 , both intersected with the unit ball

the exponential measure on \mathbb{R}_+^2 with moments defined in (2.7). With $f_{\mathbf{A}}(\mathbf{x}) := ax_1^2 + 2bx_1x_2 + cx_2^2$ and $d = 1$, the condition $\mathbf{M}_d(f_{\mathbf{A}} \mathbf{y}) \succeq 0$ which reads

$$(2.15) \quad 2 \begin{bmatrix} a + b + c & 3a + 2b + c & a + 2b + 3c \\ 3a + 2b + c & 12a + 6b + 2c & 3a + 4b + 3c \\ a + 2b + 3c & 3a + 4b + 3c & 2a + 6b + 12c \end{bmatrix} \succeq 0,$$

defines the convex cone $\mathcal{C}_1 \subset \mathbb{R}^3$. It is a connected component of the basic semi-algebraic set $\{(a, b, c) : \det(\mathbf{M}_1(f_{\mathbf{A}} \mathbf{y})) \geq 0\}$, that is, elements (a, b, c) such that:

$$(2.16) \quad 3a^3 + 15a^2b + 29a^2c + 16ab^2 + 50abc + 29ac^2 + 4b^3 + 16b^2c + 15bc^2 + 3c^3 \geq 0.$$

Alternatively, by homogeneity, instead of μ we may take the probability measure ν uniformly supported on the simplex Δ and with moments $\tilde{\mathbf{y}} = (\tilde{y}_\alpha)$ given in (2.11), in which case the corresponding matrix $\mathbf{M}_1(f_{\mathbf{A}} \tilde{\mathbf{y}})$ now reads:

$$(2.17) \quad \frac{1}{360} \begin{bmatrix} 30(a + b + c) & 6(3a + 2b + c) & 6(a + 2b + 3c) \\ 6(3a + 2b + c) & 12a + 6b + 2c & 3a + 4b + 3c \\ 6(a + 2b + 3c) & 3a + 4b + 3c & 2a + 6b + 12c \end{bmatrix} \succeq 0.$$

It turns out that up to a multiplicative factor both matrices (2.15) and (2.17) have same determinant and so define the same convex cone \mathcal{C}_1 (the same connected component of (2.16)). Figure 1 below displays the projection on the (a, b) -plane of the sets \mathcal{C}_1 and \mathcal{C} intersected with the unit ball.

2.3. An alternative representation of the cone \mathcal{C}_d^* . From its definition (2.13) in Theorem 2.2, the cone $\mathcal{C}_d^* \subset \mathcal{S}^n$ is defined through the matrix variable $\mathbf{X} \in \mathcal{S}^{s(d)}$ which lives in a (lifted) space of dimension $s(d)(s(d) + 1)/2$ and with the linear

matrix inequality (LMI) constraint $\mathbf{X} \succeq 0$ of size $s(d)$. In contrast, the convex cone $\mathcal{C}_d \subset \mathcal{S}^n$ defined in (2.10) is defined solely in terms of the entries of the matrix $\mathbf{A} \in \mathcal{S}^n$, that is, no projection from a higher dimensional space (or no lifting) is needed.

We next provide another explicit description on how \mathcal{C}_d^* can be generated with no LMI constraint and with only $s(d)$ variables, but of course this characterization is not well suited for optimization purposes.

Since every $g \in \Sigma[\mathbf{x}]_d$ can be written $\sum_{\ell} g_{\ell}^2$ for finitely many polynomials $(g_{\ell}) \subset \mathbb{R}[\mathbf{x}]_d$, the convex cone $\Sigma[\mathbf{x}]_d$ of s.o.s. polynomials can be written

$$\Sigma[\mathbf{x}]_d = \text{conv} \{g^2 : g \in \mathbb{R}[\mathbf{x}]_d\}.$$

Next, for $g \in \mathbb{R}[\mathbf{x}]_d$, with vector of coefficients $\mathbf{g} = (g_{\alpha}) \in \mathbb{R}^{s(d)}$, let $\mathbf{g}^{(2)} = (g_{\alpha}^{(2)}) \in \mathbb{R}^{s(2d)}$ be the vector of coefficients of g^2 , that is,

$$g(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_d^n} g_{\alpha} \mathbf{x}^{\alpha} \quad \rightarrow \quad g(\mathbf{x})^2 = \sum_{\alpha \in \mathbb{N}_{2d}^n} g_{\alpha}^{(2)} \mathbf{x}^{\alpha}.$$

Notice that for each $\alpha \in \mathbb{N}_{2d}^n$, $g_{\alpha}^{(2)}$ is quadratic in \mathbf{g} . For instance, with $n = 2$ and $d = 1$, $g(\mathbf{x}) = g_{00} + g_{10}x_1 + g_{01}x_2$ with $\mathbf{g} = (g_{00}, g_{10}, g_{01})^T \in \mathbb{R}^{s(1)}$, and so

$$\mathbf{g}^{(2)} = (g_{00}^2, 2g_{00}g_{10}, 2g_{00}g_{01}, g_{10}^2, 2g_{10}g_{01}, g_{01}^2)^T \in \mathbb{R}^{s(2)}.$$

Next, for $g \in \mathbb{R}[\mathbf{x}]_d$ and every $1 \leq i \leq j \leq n$, let $\mathbf{G}_d = (G_{ij}) \in \mathcal{S}^n$ be defined by:

$$(2.18) \quad G_{ij} := \int_{\mathbb{R}^n} g(\mathbf{x})^2 x_i x_j d\mu(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_{2d}^n} g_{\alpha}^{(2)} (\alpha_i + 1)! (\alpha_j + 1)! \prod_{k \neq i, j} \alpha_k!,$$

for all $1 \leq i \leq j \leq n$. We can now describe \mathcal{C}_d^* .

Corollary 2.4. *Let $\mathbf{G}_d \in \mathcal{S}^n$ be as in (2.18). Then:*

$$(2.19) \quad \mathcal{C}_d^* = \text{cl} \left(\text{conv} \{ \mathbf{G}_d : \mathbf{g} \in \mathbb{R}^{s(d)} \} \right).$$

Proof. From Theorem 2.2

$$\mathcal{C}_d^* = \text{cl} \left\{ \left(\langle \mathbf{X}, \mathbf{M}_d(x_i x_j \mathbf{y}) \rangle \right)_{1 \leq i \leq j \leq n} : \mathbf{X} \in \mathcal{S}_+^{s(d)} \right\}.$$

As in the proof of Theorem 2.2, writing $\mathbf{X} \succeq 0$ as $\sum_{k=1}^s \mathbf{g}_k \mathbf{g}_k^T$ for some vectors $(\mathbf{g}_k) \subset \mathbb{R}^{s(d)}$ (and associated polynomials $(g_k) \subset \mathbb{R}[\mathbf{x}]_d$),

$$\langle \mathbf{X}, \mathbf{M}_d(x_i x_j \mathbf{y}) \rangle = \sum_{k=1}^s \int_{\mathbb{R}_+^n} x_i x_j g_k(\mathbf{x})^2 d\mu = \sum_{k=1}^s G_{ij}^k,$$

for all $1 \leq i \leq j \leq n$, where G_{ij}^k is as in (2.18) (but now associated with g_k instead of g). Hence

$$\left(\langle \mathbf{X}, \mathbf{M}_d(x_i x_j \mathbf{y}) \rangle \right)_{1 \leq i \leq j \leq n} = \sum_{k=1}^s \mathbf{G}_d^k \in \text{conv} \{ \mathbf{G}_d : \mathbf{g} \in \mathbb{R}^{s(d)} \}.$$

Hence $\mathcal{C}^* \subseteq \text{cl} \left(\text{conv} \{ \mathbf{G}_d : \mathbf{g} \in \mathbb{R}^{s(d)} \} \right)$.

Conversely, let $h \in \text{conv} \{ \mathbf{G}_d : \mathbf{g} \in \mathbb{R}^{s(d)} \}$, i.e., for some integer s and polynomials $(g_k) \subset \mathbb{R}[\mathbf{x}]_d$, and for all $1 \leq i \leq j \leq n$,

$$\begin{aligned} h_{ij} &= \sum_{k=1}^s \underbrace{\lambda_k}_{>0} G_{ij}^k = \int_{\mathbb{R}_+^n} \left(\sum_{k=1}^s \lambda_k g_k^2(\mathbf{x}) \right) x_i x_j d\mu(\mathbf{x}) \\ &= \langle \mathbf{X}, \mathbf{M}_d(x_i x_j \mathbf{y}) \rangle \end{aligned}$$

where $\mathbf{X} = \sum_{k=1}^s \lambda_k \mathbf{g}_k \mathbf{g}_k^T \in \mathcal{S}_+^{s(d)}$. Hence $\text{cl}(\text{conv} \{ \mathbf{G}_d : \mathbf{g} \in \mathbb{R}^{s(d)} \}) \subseteq \mathcal{C}^*$. \square

The characterization (2.19) of \mathcal{C}_d^* should be compared with the characterization $\text{conv} \{ \mathbf{x}\mathbf{x}^T : \mathbf{x} \in \mathbb{R}_+^n \}$ of \mathcal{C}^* .

Example 2. With $n = 2$ and $d = 1$, \mathbf{G}_1 reads:

$$\begin{bmatrix} 2g_{00}^2 + 12g_{10}(g_{00} + g_{01}) + 4g_{01}(g_{00} + g_{01}) + 24g_{10}^2 & g_{00}^2 + 4g_{00}(g_{10} + g_{01}) + 6(g_{10}^2 + g_{01}^2) + 8g_{10}g_{01} \\ g_{00}^2 + 4g_{00}(g_{10} + g_{01}) + 6(g_{10}^2 + g_{01}^2) + 8g_{10}g_{01} & 2g_{00}^2 + 4g_{10}(g_{00} + g_{10}) + 12g_{01}(g_{00} + g_{10}) + 24g_{01}^2 \end{bmatrix}$$

2.4. \mathcal{K} -copositive and \mathcal{K} -completely positive matrices. Let $\mathcal{K} \subset \mathbb{R}^n$ be a closed convex cone and let $\mathcal{C}_{\mathcal{K}}$ be the convex cone of \mathcal{K} -copositive matrices, i.e., matrices $\mathbf{A} \in \mathcal{S}^n$ such that $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ on \mathcal{K} . Its dual cone $\mathcal{C}_{\mathcal{K}}^* \subset \mathcal{S}^n$ is the cone of \mathcal{K} -completely positive matrices.

Then one may define a hierarchy of convex cones $(\mathcal{C}_{\mathcal{K}})_d$ and $(\mathcal{C}_{\mathcal{K}}^*)_d$, $d \in \mathbb{N}$, formally exactly as in Theorem 2.1 and Theorem 2.2, but now \mathbf{y} is the moment sequence of a finite Borel measure μ with $\text{supp } \mu = \mathcal{K}$ (instead of $\text{supp } \mu = \mathbb{R}_+^n$ in (2.7)), i.e., $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, with:

$$(2.20) \quad y_\alpha := \int_{\mathcal{K}} \mathbf{x}^\alpha d\mu, \quad \forall \alpha \in \mathbb{N}^n \quad (\text{and where } \text{supp } \mu = \mathcal{K}).$$

And so with \mathbf{y} as in (2.20) Theorem 2.1 and 2.2, as well as Corollary 2.3, are still valid. (In (2.14) replace $\int_{\mathbb{R}_+^n}$ with $\int_{\mathcal{K}}$). Therefore,

$$\mathcal{C}_{\mathcal{K}} = \bigcap_{d=0}^{\infty} (\mathcal{C}_{\mathcal{K}})_d \quad \text{and} \quad \mathcal{C}_{\mathcal{K}}^* = \text{cl} \bigcup_{d=0}^{\infty} (\mathcal{C}_{\mathcal{K}}^*)_d.$$

But of course, for practical implementation, one need to know the sequence $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, which was easy when $\mathcal{K} = \mathbb{R}_+^n$. For instance, if \mathcal{K} is a polyhedral cone, by homogeneity of $f_{\mathbf{A}}$, one may equivalently consider a compact base of \mathcal{K} (which is a polytope \mathcal{K}'), and take for μ the Lebesgue measure on \mathcal{K}' . Then all moments of μ can be calculated exactly. The same argument works for every convex cone \mathcal{K} for which one may compute all moments of a finite Borel measure μ whose support is a compact base of \mathcal{K} .

ACKNOWLEDGEMENT

The author wishes to thank two anonymous referees for their very helpful remarks and suggestions to improve the initial version of this note.

REFERENCES

- [1] K.M. Anstreicher, S. Burer. Computable representations for convex hulls of low-dimensional quadratic forms, Math. Program. Sér. B **124** (2010), pp. 33–43.
- [2] I.M. Bomze. Copositive optimization - recent developments and applications, European J. Oper. Res., 2011. To appear.

- [3] I.M. Bomze, W. Schachinger and G. Uchida. Think co(mpletely) positive! - matrix properties, examples and a clustered bibliography on copositive optimization, *J. Global Optim.* 2011. To appear.
- [4] I.M. Bomze, E. de Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming, *J. Global Optim.* **24** (2002), 163–185.
- [5] S. Burer. Copositive programming, in *Handbook of Semidefinite, Conic and Polynomial Optimization*, M. Anjos and J.B. Lasserre, Eds., Springer, New York, 2012, pp. 201–218.
- [6] E. de Klerk, D.V. Pasechnik. A linear programming formulation of the standard quadratic optimization problem, *J. Global Optim.* **37** (2007), pp. 75–84.
- [7] E. de Klerk, D.V. Pasechnik. Approximation of the stability number of a graph via copositive programming, *SIAM J. Optim.* **12** (2002), 875–892.
- [8] M. Dür. Copositive Programming - a survey, In *Recent Advances in Optimization and its Applications in Engineering*, M. Diehl, F. Glineur, E. Jarlebring and W. Michiels Eds., Springer, New York, 2010, pp. 3 – 20.
- [9] J.W. Gaddum. Linear inequalities and quadratic forms, *Pacific J. Math.* **8** (1958), pp. 411–414.
- [10] A. Grundmann and H.M. Moeller. Invariant integration formulas for the n-simplex by combinatorial methods, *SIAM J. Numer. Anal.* **15** (1978), pp. 282–290.
- [11] J.-B. Hiriart-Urruty, A. Seeger. A variational approach to copositive matrices, *SIAM Rev.* **52** (2010), 593–629.
- [12] J.B. Lasserre. A new look at nonnegativity and polynomial optimization, *SIAM J. Optim.* **21**(2011), pp. 864–885.
- [13] P. Parrilo. Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization, Ph.D. Dissertation, California Institute of Technology, 2000.
- [14] J. Peña, J. Vera, L. Zuluaga. Computing the stability number of a graph via linear and semidefinite programming, *SIAM J. Optim.* **18** (2007), 87–105.
- [15] R.T. Rockafellar. *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
- [16] K. Schmüdgen. The K -moment problem for compact semi-algebraic sets, *Math. Ann.* **289** (1991), 203–206.
- [17] M. Schweighofer. Global optimization of polynomials using gradient tentacles and sums of squares, *SIAM J. Optim.* **17** (2006), pp. 920–942.
- [18] G. Stengel. A Nullstellensatz and a Positivstellensatz in semialgebraic geometry, *Math. Ann.* **207**, pp. 87–97.
- [19] L. Vandenberghe and S. Boyd. Semidefinite programming, *SIAM Rev.* **38** (1996), pp. 49–95.
- [20] E. A. Yildirim. On the accuracy of uniform polyhedral approximations of the copositive cone, *Optim. methods Softw.*, 2011. To appear.

LAAS-CNRS AND INSTITUTE OF MATHEMATICS, UNIVERSITY OF TOULOUSE, LAAS, 7 AVENUE DU COLONEL ROCHE, 31077 TOULOUSE CÉDEX 4, FRANCE
E-mail address: `lasserre@laas.fr`