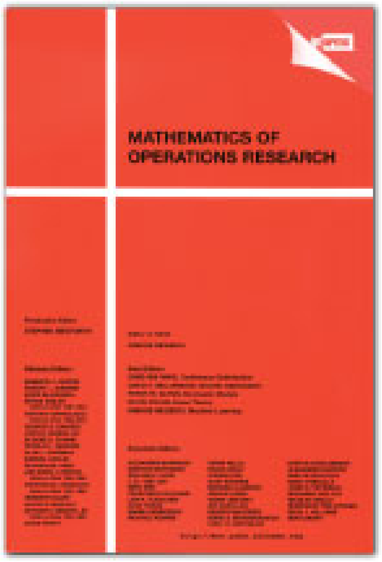

Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

On Cones of Nonnegative Quadratic Functions

Jos F. Sturm, Shuzhong Zhang,

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 28, No. 2, May 2003, pp. 246–267
Printed in U.S.A.

ON CONES OF NONNEGATIVE QUADRATIC FUNCTIONS

JOS F. STURM AND SHUZHONG ZHANG

We derive linear matrix inequality (LMI) characterizations and dual decomposition algorithms for certain matrix cones which are generated by a given set using generalized co-positivity. These matrix cones are in fact cones of nonconvex quadratic functions that are nonnegative on a certain domain. As a domain, we consider for instance the intersection of a (upper) level-set of a quadratic function and a half-plane. Consequently, we arrive at a generalization of Yakubovich's S-procedure result. Although the primary concern of this paper is to characterize the matrix cones by LMIs, we show, as an application of our results, that optimizing a general quadratic function over the intersection of an ellipsoid and a half-plane can be formulated as semidefinite programming (SDP), thus proving the polynomiality of this class of optimization problems, which arise, e.g., from the application of the trust region method for nonlinear programming. Other applications are in control theory and robust optimization.

1. Introduction. In mathematics it is important to study functionals that are nonnegative over a given domain. As an example, the concept of duality is based on such a consideration and in convex analysis, the dual (polar) of a cone consists exactly of all the linear mappings that are nonnegative (nonpositive) over the cone itself. As another example, the positive semidefinite matrices are defined as the quadratic forms that are nonnegative over the whole Euclidean space. No doubt these are extremely important concepts. Recently, optimization with positive semidefiniteness restrictions (linear matrix inequalities), known as semidefinite programming, or SDP for short, received a lot of attention (see Wolkowicz et al. 2000 and references therein).

In this paper, we shall apply the power of SDP to solve problems involving general quadratic functions. We first introduce the cones formed by quadratic functions that are nonnegative over a given region. Properties of such cones are discussed. In some special cases, we are able to characterize these cones using linear matrix inequalities (LMIs). The characterization leads us to solve several new classes of optimization problems, arising e.g. from the trust region method for nonlinear programming (Rendl and Wolkowicz 1997, Fu et al. 1998). The results also provide new tools for robust optimization (El Ghaoui and Lebret 1997, Ben-Tal and Nemirovsky 1998), in which the constraints can now depend in a quadratic fashion on the uncertain parameter.

Our results can be considered as extensions of Yakubovich's S-procedure result (Yakubovich 1971, Fradkov and Yakubovich 1973), also known as the S-lemma, which characterizes quadratic functions that are nonnegative over the domain defined by another single quadratic function. Our extension characterizes the cone of quadratic functions that are nonnegative over a domain that is defined by two convex quadratic constraints, one of them being affine. The characterization does not fall in the S-procedure class, but is a mixture of a semidefinite, a second-order cone, and a linear constraint.

Unlike our approach, the original proofs of Yakubovich (1971) and Fradkov and Yakubovich (1973) were based on separation and range convexity results. The classical

Received April 14, 2001; revised March 24, 2002.
MSC 2000 subject classification. Primary: 90C22, 90C20.
OR/MS subject classification. Primary: Programming/quadratic.
Key words. LMI, SDP, co-positive cones, quadratic functions, S-procedure, matrix decomposition.

0364-765X/03/2802/0246/$05.00
1526-5471 electronic ISSN, © 2003, INFORMS

range convexity result by Dines (1941) states that for arbitrary $n \times n$ real symmetric matrices F and G, the set

$$\left\{ \begin{bmatrix} x^T F x \\ x^T G x \end{bmatrix} \,\middle|\, x \in \Re^n \right\} \subseteq \Re^2$$

is convex. Fradkov and Yakubovich (1973) further showed that convexity in $\Re^3$ holds with three Hermitian quadratic forms, and x running over all complex n-tuples instead of merely the real n-tuples. Another extension, due to Barvinok (1975), is as follows. Let $F_1, F_2, \ldots, F_m$ be real symmetric matrices. Then the set

$$\left\{ y \in \Re^m \mid y_i = \operatorname{tr} X^T F_i X,\ i = 1, \ldots, m,\ X \in \Re^{n \times r} \right\}$$

is convex whenever r is at least the integer part of $((\sqrt{8m+1} - 1)/2)$. As recently shown by Hiriart-Urruty and Torki (2001), these results can be explained from the dimension of the faces of the cone of (real symmetric or complex Hermitian) positive semidefinite matrices. An interesting relationship between the rank of an extremal matrix and the dimension of the faces of a LMI is given in Theorem 2.1 in Pataki (1998). Range convexity of general nonlinear maps where x runs over a small ball in $\Re^n$ has recently been proved by Polyak (2001).

An important concept that is used in our approach is "co-positivity over a domain D," which reduces to the usual concept of co-positivity when D is the nonnegative orthant (i.e., $D = \Re^n_+$). When D is a polyhedral cone, we arrive at the generalized co-positivity concept of Quist et al. (1998). Recent results on approximating the usual co-positive cone by SDP can be found in Parrilo (2000) and De Klerk and Pasechnik (2000).

We shall show that the concept of co-positivity over a domain D is dual to the concept of decomposibility of a positive semidefinite matrix into rank-1 matrices $x_i x_i^T$ with $x_i \in D$; see Proposition 1 in this paper. We have developed efficient decomposition algorithms for the tractable dual cones that are considered in this paper. The dual matrix corresponding to a quadratic function that is designed in the primal can thus be decomposed into feasible roots of that quadratic function. The rank of the dual matrix corresponds to the number of feasible roots. Our novel dual approach to S-lemma type results is, hence, useful for efficiently computing roots from dual matrix solutions.

References on quadratic systems and error bounds can be found in Luo and Sturm (2000). Some recent results on LMIs and nonnegativity expressed as *sum of squares* (SOS) can be found in Parrilo (2000) and Nesterov (2000). One may consider the SOS approach as "primal," whereas the approach in our paper is "dual." A general dual approach to global optimization with polynomials based on the problem of moments has recently been proposed by Lasserre (2001).

The organization of this paper is as follows. We introduce our definitions and notation concerning co-positivity with respect to a cone, cones of nonnegative quadratic functions on a specified domain, as well as the concept of homogenization in §2. Section 3 is devoted to a possible application of our analysis, namely, nonconvex quadratic optimization. We describe how general nonconvex quadratic optimization problems can be reformulated as conic linear programming over cones of nonnegative quadratic functions. In §4 we investigate the cones that are obtained by homogenization of a domain that is given as the intersection of upper level sets of some quadratic functions. Then, in §5, two matrix decomposition results are proved in a constructive way. The results serve the purpose of characterizing, in terms of LMIs, cones of nonnegative quadratic functions for three different classes of domains of nonnegativity. The domains considered are defined either by a nonconvex quadratic inequality, by an equality constraint in a strictly concave (or convex) quadratic function, or by the combination of a convex quadratic inequality and a linear (affine) inequality. Based on the technique of SDP, these results imply among others the polynomial solvability of nonconvex quadratic optimization problems over (unions of) these three classes of domains. We conclude the paper in §6. We want to remark that the material of §3 is merely an illustration, and the reader can skip this section if desired. After reading §2, it is possible to proceed immediately with §5, which includes our main results, and track back to the technical lemmas in §4 whenever they are referred to.

NOTATION. Given a set D in a Euclidean space, we let cone(D) denote the convex cone consisting of all nonnegative combinations of elements of D. Similarly, we let conv(D) denote the convex set consisting of all convex combinations of elements of D. If D is a cone, then $\text{conv}(D) = \text{cone}(D)$. We associate with a cone K in a Euclidean space the dual cone $K^* := \{y \mid x \bullet y \geq 0 \ \forall x \in K\}$, where $\bullet$ denotes the standard inner product of the Euclidean space. We let $\mathcal{S}^{n\times n}$ denote the $n(n+1)/2$-dimensional Euclidean space of symmetric $n \times n$ matrices, with the standard inner product

$$X \bullet Y = \text{ tr } XY = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}y_{ij},$$

for $X, Y \in \mathcal{S}^{n\times n}$. We let $\mathcal{S}_+^{n\times n}$ denote the cone of positive semidefinite matrices in $\mathcal{S}^{n\times n}$. Also, "$X \succ 0$" ("$X \succeq 0$") means that X is a symmetric positive definite (positive semidefinite) matrix.

2. Preliminaries. Let $D \subseteq \Re^n$ be a given set. Consider all symmetric matrices that are co-positive over D, i.e.

$$\mathcal{C}_+(D) := \{Z \in \mathcal{S}^{n\times n} \mid x^T Z x \geq 0, \ \forall x \in D\}. \tag{1}$$

It is obvious that $\mathcal{C}_+(D)$ is a closed convex cone, and that

$$\mathcal{C}_+(D) = \mathcal{C}_+(D \cup (-D)). \tag{2}$$

We also have an obvious dual characterization of $\mathcal{C}_+(D)$, namely:

PROPOSITION 1. *It holds that*

$$\mathcal{C}_+(D) = \left(\text{cone}\left\{yy^T \mid y \in D\right\}\right)^*.$$

PROOF. If $X \in \mathcal{C}_+(D)$, then by definition $0 \leq y^T X y = X \bullet (yy^T)$ for all $y \in D$. Because the sum of nonnegative quantities is nonnegative, it follows that $X \bullet Z \geq 0$ whenever Z is a nonnegative combination of matrices in $\{yy^T \mid y \in D\}$. This establishes $\mathcal{C}_+(D) \subseteq (\text{cone}\{yy^T \mid y \in D\})^*$. Conversely, if $X \bullet Z \geq 0$ for all $Z \in \text{cone}\{yy^T \mid y \in D\}$, then certainly $0 \leq X \bullet (yy^T) = y^T X y$ for all $y \in D$ and, hence, $X \in \mathcal{C}_+(D)$. □

Clearly, $\mathcal{C}_+(\Re^n) = \mathcal{S}_+^{n\times n}$ is the set of positive semidefinite matrices. In another well-known case, where $D = \Re_+^n$, the set $\mathcal{C}_+(D)$ is called the *co-positive cone*. Testing whether a given matrix belongs to the co-positive cone is co-NP-hard, i.e., testing whether it does not belong to the co-positive cone is NP-hard (see Murty and Kabadi 1987). We remark for general D that the validity of the claim "$Z \notin \mathcal{C}_+(D)$" can be certified by a vector "$x \in D$" for which $x^T Z x < 0$; this decision problem is therefore in NP, provided that $x \in D$ is easy to check.

Two classical theorems from convex analysis are particularly worth mentioning in the context of this paper: the bi-polar theorem and Carathéodory's theorem (Rockafellar 1970, Carathéodory 1907). The *bi-polar theorem* states that if $K \subseteq \Re^n$ is a convex cone, then $(K^*)^* = \text{cl}(K)$, i.e., dualizing K twice yields the closure of K. *Carathéodory's theorem* states that for any set $S \subseteq \Re^n$ it holds that $x \in \text{conv}(S)$ if and only if there exist $y_1, y_2, \ldots, y_{n+1}$ such that $x = \sum_{i=1}^{n+1} \alpha_i y_i$ for some $\alpha_i \geq 0$ with $\sum_{i=1}^{n+1} \alpha_i = 1$.

Using the bipolar theorem, it follows from Proposition 1 that $\mathscr{C}_+(D)^* = \operatorname{cl}\operatorname{cone}\{yy^T \mid y \in D\}$. The following lemma, which is based on Carathéodory's theorem, implies further that $\mathscr{C}_+(D)^* = \operatorname{cone}\{yy^T \mid y \in \operatorname{cl}(D)\}$.

LEMMA 1. *Let $D \subseteq \Re^n$. Then,*

$$\operatorname{cl}\operatorname{cone}\{yy^T \mid y \in D\} = \operatorname{cone}\{yy^T \mid y \in \operatorname{cl}(D)\}.$$

PROOF. Suppose that $Z \in \operatorname{cl}\operatorname{cone}\{yy^T \mid y \in D\}$, then $Z = \lim_{k\to\infty} Z_k$ for some $Z_k \in \operatorname{cone}\{yy^T \mid y \in D\}$. Because the dimension of $\mathscr{S}^{n\times n}$ is $N := n(n+1)/2$, it follows from Carathéodory's theorem that for given Z_k there exists an $n \times (N+1)$ matrix Y_k such that $Z_k = Y_k Y_k^T$, and each column of Y_k is a positive multiple of a vector in D. Furthermore, we have

$$\|Y_k\|_F^2 = \operatorname{tr} Y_k Y_k^T = \operatorname{tr} Z_k \to \operatorname{tr} Z.$$

Therefore, the sequence $Y_1, Y_2, \ldots$ is bounded, and must have a cluster point Y^* for $k \to \infty$. Obviously, each column of Y^* is then a positive multiple of a vector in $\operatorname{cl}(D)$, and because $Z = Y^*(Y^*)^T$, it follows that $Z \in \operatorname{cone}\{yy^T \mid y \in \operatorname{cl}(D)\}$. The converse relationship is trivial. $\square$

By definition, $\mathscr{C}_+(D)$ consists of all quadratic forms that are nonnegative on D. We shall now consider the cone of all nonnegative quadratic functions (not necessarily homogeneous) that are nonnegative on D. Namely, we define

$$\mathscr{F}\mathscr{C}_+(D) := \left\{ \begin{bmatrix} z_0, & z^T \\ z, & Z \end{bmatrix} \,\middle|\, z_0 + 2z^T x + x^T Z x \geq 0,\ \forall x \in D \right\}. \tag{3}$$

For a quadratic function $q(x) = c + 2b^T x + x^T A x$, we introduce its matrix representation, denoted by

$$\mathrm{M}(q(\cdot)) = \begin{bmatrix} c, & b^T \\ b, & A \end{bmatrix}. \tag{4}$$

In this notation, $q(x) \geq 0$ for all $x \in D$ if and only if $\mathrm{M}(q(\cdot)) \in \mathscr{F}\mathscr{C}_+(D)$.

To derive a dual characterization of the matrix cone $\mathscr{F}\mathscr{C}_+(D)$, we need the concept of *homogenization*. Formally, for a set D, its homogenization is given by

$$\mathscr{H}(D) = \operatorname{cl}\left\{ \begin{bmatrix} t \\ x \end{bmatrix} \in \Re_{++} \times \Re^n \,\middle|\, x/t \in D \right\},$$

which is a closed cone (not necessarily convex). If D is a bounded set, then

$$\mathscr{H}(D) = \left\{ \begin{bmatrix} t \\ x \end{bmatrix} \,\middle|\, t > 0,\ x/t \in \operatorname{cl}(D) \right\} \cup \{0\}. \tag{5}$$

Otherwise, this may not be true. A simple example is $D = [1, +\infty)$; in this case $[0,\ 1]^T \in \mathscr{H}(D)$. As another example, $\mathscr{H}(\Re^n) = \Re_+ \times \Re^n$. The following proposition states that the nonnegative quadratic functions on D and the nonnegative quadratic forms on $\mathscr{H}(D)$ are the same geometric objects, hence our interest in the concept of homogenization.

PROPOSITION 2. *For any set $D \neq \varnothing$, it holds that*

$$\mathscr{F}\mathscr{C}_+(D) = \mathscr{C}_+(\mathscr{H}(D)) = \mathscr{C}_+\left(\mathscr{H}(D) \cup (-\mathscr{H}(D))\right).$$

PROOF. The second identity is a special case of relation (2). Furthermore, to see that $\mathcal{C}_+(\mathcal{H}(D)) \subseteq \mathcal{F}\mathcal{C}_+(D)$, it suffices to observe that $x \in D$ implies $[1, \; x^T]^T \in \mathcal{H}(D)$ by definition of $\mathcal{H}(D)$. It remains to show that $\mathcal{F}\mathcal{C}_+(D) \subseteq \mathcal{C}_+(\mathcal{H}(D))$.

Let $[t, \; x^T]^T \in \mathcal{H}(D)$, i.e. there exist $t_k > 0$ and $x_k/t_k \in D$ such that $t = \lim_{k\to\infty} t_k$ and $x = \lim_{k\to\infty} x_k$. Any

$$\begin{bmatrix} z_0, & z^T \\ z, & Z \end{bmatrix} \in \mathcal{F}\mathcal{C}_+(D)$$

necessarily satisfies

$$z_0 + 2z^T(x_k/t_k) + (x_k/t_k)^T Z (x_k/t_k) \geq 0,$$

or equivalently

$$z_0 t_k^2 + 2t_k z^T x_k + x_k^T Z x_k \geq 0.$$

By taking limits we get

$$z_0 t^2 + 2tz^T x + x^T Z x \geq 0,$$

which leads to the conclusion that

$$\begin{bmatrix} z_0, & z^T \\ z, & Z \end{bmatrix} \in \mathcal{C}_+(\mathcal{H}(D)). \quad \square$$

Combining Proposition 2 with Proposition 1, we arrive at the following corollary.

COROLLARY 1. *For any nonempty set D, it holds that*

$$\mathcal{F}\mathcal{C}_+(D) = \operatorname{conv}\{yy^T \mid y \in \mathcal{H}(D)\}^*.$$

Using Lemma 1 and the fact that $\mathcal{H}(D)$ is, by definition, a closed cone, we can dualize Corollary 1 to

$$\mathcal{F}\mathcal{C}_+(D)^* = \operatorname{conv}\{yy^T \mid y \in \mathcal{H}(D)\}. \tag{6}$$

We remark from Proposition 2 that

$$\mathcal{F}\mathcal{C}_+(\Re^n) = \mathcal{C}_+\left(\mathcal{H}(\Re^n) \cup (-\mathcal{H}(\Re^n))\right) = \mathcal{C}_+(\Re^{n+1}) = \mathcal{S}_+^{(1+n)\times(1+n)}. \tag{7}$$

In other words, the cone of $(n+1)\times(n+1)$ positive semidefinite matrices is equal to the cone of (matrix representations of) quadratic functions that are nonnegative on the entire domain $\Re^n$.

Another case that deserves special attention is the sphere with radius 1 centered at the origin,

$$B(n) := \left\{ x \in \Re^n \mid \|x\| \leq 1 \right\}.$$

Because this is a bounded set, we may apply (5) to conclude that

$$\mathcal{H}(B(n)) = \left\{ \begin{bmatrix} t \\ x \end{bmatrix} \middle| \; \|x\| \leq t \right\} =: \mathrm{SOC}(n+1).$$

We see that the homogenization of $B(n)$ is the Lorentz cone, or second-order cone, denoted by $\mathrm{SOC}(n+1)$. According to Corollary 1, it holds that

$$\mathcal{F}\mathcal{C}_+(B(n)) = (\operatorname{conv}\{yy^T \mid y \in \mathrm{SOC}(n+1)\})^*.$$

In §4, we will consider (among others) domains of the form $D = \{x \mid q(x) \geq 0\}$, where $q(\cdot)$ is a given quadratic function. Choosing $q(x) = 1 - x^T x$ yields $D = B(n)$. In §5, we will see as a special case of Theorem 1 that

$$\mathrm{conv}\{yy^T \mid y \in \mathrm{SOC}(n+1)\} = \{X \in \mathcal{S}^{(1+n)\times(1+n)} \mid X \succeq 0,\ J \bullet X \geq 0\},$$

where $J := \mathrm{M}(q(\cdot))$, i.e.

$$J = \begin{bmatrix} 1, & 0 \\ 0, & -I \end{bmatrix} \in \mathcal{S}^{(1+n)\times(1+n)}.$$

This will then easily lead to the relation

$$\mathcal{F}\mathcal{C}_+(B(n)) = \{Z \mid Z - tJ \succeq 0,\ t \geq 0\},$$

which is known from Rendl and Wolkowicz (1997) and Fu et al. (1998).

3. Global nonconvex quadratic optimization. Consider the general nonconvex quadratic optimization problem

$$\text{(P)} \quad \inf\{f(x) \mid x \in D\},$$

where $f(\cdot)$ is a (nonconvex) quadratic function and $D \subset \Re^n$ is a possibly nonconvex domain. Let N be an arbitrary positive integer. Then,

$$\inf\{f(x) \mid x \in D\} = \inf\left\{ \sum_{i=1}^N t_i^2 f(x_i) \,\middle|\, \sum_{j=1}^N t_j^2 = 1 \text{ and } x_i \in D,\ i = 1, 2, \ldots, N \right\}.$$

Namely, $f(x)$ with $x \in D$ can never be smaller than the right-hand side, because one may set $x_i = x$ for all i. Conversely, $\sum_{i=1}^N t_i^2 f(x_i)$ can never be smaller than the left-hand side because

$$\sum_{i=1}^N t_i^2 f(x_i) \geq \min_{j=1,2,\ldots,N}\{f(x_j)\} \cdot \left(\sum_{i=1}^N t_i^2\right) = \min_{j=1,2,\ldots,N}\{f(x_j)\} \geq \inf\{f(x) \mid x \in D\}.$$

Using the matrix representation of $f(\cdot)$, we have

$$t_i^2 f(x_i) = t_i^2 \begin{bmatrix} 1 \\ x_i \end{bmatrix}^T \mathrm{M}(f(\cdot)) \begin{bmatrix} 1 \\ x_i \end{bmatrix} = y_i^T\ \mathrm{M}(f(\cdot)) y_i, \quad y_i := \begin{bmatrix} |t_i| \\ |t_i| x_i \end{bmatrix}.$$

Obviously, $x_i \in D$ implies $y_i \in \mathcal{H}(D)$. Conversely, we have for any $y = [t,\ \xi^T]^T \in \mathcal{H}(D)$ with $e_1^T y = t > 0$, where e_1 denotes the first column of the identity matrix, that

$$y^T\ \mathrm{M}(f(\cdot)) y = t^2 f(\xi/t) \geq t^2 \inf\{f(x) \mid x \in D\}.$$

By definition of $\mathcal{H}(D)$, it thus follows that if $\inf\{f(x) \mid x \in D\} > -\infty$, then

$$\text{(8)} \qquad y \in \mathcal{H}(D) \quad \Rightarrow \quad y^T\ \mathrm{M}(f(\cdot)) y \geq (e_1^T y)^2 \inf\{f(x) \mid x \in D\}.$$

Therefore,

$$\inf\{f(x) \mid x \in D\} = \inf\left\{ \sum_{i=1}^N y_i^T\ \mathrm{M}(f(\cdot)) y_i \,\middle|\, \sum_{j=1}^N (e_1^T y_j)^2 = 1 \text{ and } y_i \in \mathcal{H}(D),\ i = 1, 2, \ldots, N \right\}.$$

Because the above relation holds in particular for $N = 1 + n(n+1)/2$, it follows from Carathéodory's theorem that

$$\inf\{f(x) \mid x \in D\} = \inf\left\{ \mathrm{M}(f(\cdot)) \bullet Z \,\middle|\, z_{11} = 1 \text{ and } Z \in \operatorname{conv}\{yy^T \mid y \in \mathscr{H}(D)\}\right\},$$

where $z_{11} = e_1^T Z e_1$ denotes the $(1,1)$-entry of Z. Using also (6), we conclude that the nonconvex problem (P) is equivalent to the convex problem (MP), defined as

$$\text{(MP)} \quad \inf\{\mathrm{M}(f(\cdot)) \bullet Z \mid Z \in \mathscr{FC}_+(D)^*,\ z_{11} = 1\}.$$

Notice that if $D \neq \varnothing$, then $z_{11} > 0$ for any Z in the relative interior of $\mathscr{FC}_+(D)^*$. It follows that (MP) has a feasible solution in the relative interior of $\mathscr{FC}_+(D)^*$. Hence, (MP) satisfies the relative Slater condition, or interior point condition, which implies that there can be no duality gap, and that either (MP) is unbounded or the dual optimal value is attained (Sturm 2000). Indeed, the dual of (MP) is (MD):

$$\text{(MD)} \quad \sup\{\phi \mid \mathrm{M}(f(\cdot)) - \phi e_1 e_1^T \in \mathscr{FC}_+(D)\}.$$

Because $\mathrm{M}(f(\cdot)) - \phi e_1 e_1^T = \mathrm{M}(f(\cdot) - \phi)$, we may rewrite (MD) as

$$\sup\{\phi \mid f(x) \geq \phi \text{ for all } x \in D\},$$

and it is clear the optimal value of (MD) is indeed equal to the optimal value of (MP).

In principle, the nonconvex problem (P) and the the convex problem (MP) are completely equivalent. Namely, Carathéodory's theorem implies that if Z is a feasible solution for (MP), then there exist $y_i = [t_i,\ \ \xi_i^T]^T \in \mathscr{H}(D)$, $i = 1, 2, \ldots, N$, such that $Z = \sum_{i=1}^N y_i y_i^T$. If there is an i such that $t_i = 0$ and $y_i^T\ \mathrm{M}(f(\cdot)) y_i < 0$, then (P) must be unbounded due to (8). Otherwise, we have

$$\mathrm{M}(f(\cdot)) \bullet Z \geq \min\{f(\xi_i/t_i) \mid i \text{ such that } t_i > 0\}.$$

Equality holds if and only if

$$\begin{cases} \mathrm{M}(f(\cdot)) \bullet Z = f(\xi_i/t_i) & \text{for all } i \text{ with } t_i > 0, \\ y_i^T\ \mathrm{M}(f(\cdot)) y_i = 0 & \text{for all } i \text{ with } t_i = 0. \end{cases}$$

This shows that if Z is an optimal solution to (MP), then the decomposition $Z = \sum_{i=1}^N y_i y_i^T$ yields an optimal solution *for any* y_i with $e_1^T y_i > 0$. Because $\sum_{i=1}^N (e_1^T y_i)^2 = 1$, it yields *at least* one (global) optimal solution.

A solution to the dual problem (MD) can be used to certify *global* optimality in the primal problem (MP) or (P). We remark that the classical approach only yields *local* optimality conditions for (P). The fact that we can reformulate a general nonconvex problem (P) into a convex problem (MP) does not necessarily make such a problem easier to solve. For example, we already encountered in §2 the NP-hard problem of deciding whether a matrix is in the complement of $\mathscr{FC}_+(\Re^n_+)$. Furthermore, Carathéodory's theorem states only the existence of a decomposition of Z; it is in general not clear how such a decomposition should be constructed. Indeed it is well known that problem (P) is NP-hard (Vavavis 1991) in its general setting.

However, in all three cases that we will discuss in §5, namely,

(1) $D = \{x \mid q(x) \geq 0\}$,

(2) $D = \{x \mid q(x) = 0\}$ with $q(\cdot)$ strictly concave, and

(3) $D = \{x \mid q(x) \geq 0,\ \text{and } a^T x \geq a_0\}$ with $q(\cdot)$ concave,

the optimization problem (MP) and its dual (MD) turn out to be SDP problems, for which polynomial-time and effective solution methods exist. And furthermore, we propose efficient algorithms to decompose matrices in the dual cone $\mathcal{FC}_+(D)^*$ as a sum of rank-1 solutions in $\mathcal{FC}_+(D)^*$. Therefore, once we find a (nearly) optimal Z solution to (MP) we will also have (nearly) optimal x solutions to (P). This is remarkable, because (P) has some nasty features: the optimal solution set of (P) can be disconnected, and, in cases (1) and (2), the quadratically constrained sets "D" are not necessarily convex.

We remark that problem (MD) has only one variable and only one conic constraint. In general, however, a conic linear programming model has multiple variables and multiple conic constraints. The general framework allows for the optimal *design* of quadratic functions and for robust optimization, where the constraints depend on the uncertain parameters in a quadratic fashion. The dual matrix decomposition will then yield worst-case scenarios for the optimal robust design.

4. Quadratically constrained sets. In this section, we shall study the case when the domain D is defined by some quadratic (in)equalities. Our aim is to show that under certain conditions, $\mathcal{H}(D)$ or $\mathcal{H}(D)\cup(-\mathcal{H}(D))$ can then completely be characterized by homogeneous quadratic constraints. With such a characterization, it is then easy to check whether a given vector belongs to $\mathcal{H}(D)$; the claim that a matrix belongs to $\mathcal{FC}_+(D)^*$ can then in principle be certified due to (6) and Carathéodory's theorem.

As a first step, let us consider one quadratic function $q(x) = c + 2b^T x + x^T Ax$, and its upper level set

$$D = \{x \in \Re^n \mid q(x) \geq 0\}.$$

Obviously, $q(x) \geq 0$ for all $x \in D$, so that $\mathrm{M}(q(\cdot)) \in \mathcal{FC}_+(D)$. The following lemma characterizes the homogenized cone of D.

LEMMA 2. *Consider a quadratic function $q(x) = c + 2b^T x + x^T Ax$ for which the upper level set $D = \{x \mid q(x) \geq 0\}$ is nonempty. It holds that*

$$\mathcal{H}(D)\cup(-\mathcal{H}(D)) = \left\{ \begin{bmatrix} t \\ x \end{bmatrix} \,\middle|\, t^2 c + 2tb^T x + x^T Ax \geq 0 \right\}.$$

PROOF. We remark first that $\mathrm{M}(q(\cdot)) \in \mathcal{FC}_+(D) = \mathcal{C}_+(\mathcal{H}(D)\cup(-\mathcal{H}(D)))$, where the identity follows from Proposition 2. Therefore,

$$\begin{bmatrix} t \\ x \end{bmatrix} \in \mathcal{H}(D)\cup(-\mathcal{H}(D)) \quad \Rightarrow \quad t^2 c + 2tb^T x + x^T Ax \geq 0.$$

To show the converse, we consider a pair (t, x) in the set

$$\left\{ \begin{bmatrix} t \\ x \end{bmatrix} \,\middle|\, t^2 c + 2tb^T x + x^T Ax \geq 0 \right\}.$$

If $t > 0$, then $x/t \in D$ and so $[t,\ x^T]^T \in \mathcal{H}(D)$. If $t < 0$, then $(-x)/(-t) \in D$, and so $[t,\ x^T]^T \in -\mathcal{H}(D)$. It remains to consider the case $t = 0$. We have

$$0 \leq t^2 c + 2tb^T x + x^T Ax = x^T Ax = (-x)^T A(-x).$$

Because $D \neq \varnothing$, there must exist $\bar{x}$ such that $q(\bar{x}) \geq 0$. Let $\epsilon \in \Re \setminus \{0\}$. Then,

$$\epsilon^2 q((x+\epsilon\bar{x})/\epsilon) = \epsilon^2 q(\bar{x}) + 2\epsilon(b + A\bar{x})^T x + x^T Ax.$$

Therefore, if $(b+A\bar{x})^T x \geq 0$, then $(x+\epsilon\bar{x})/\epsilon \in D$ for all $\epsilon > 0$ and hence $[0,\ x^T]^T \in \mathcal{H}(D)$. Otherwise, i.e. $(b+A\bar{x})^T x < 0$, we have $(x+\epsilon\bar{x})/\epsilon = (-x-\epsilon\bar{x})/(-\epsilon) \in D$ for all $\epsilon < 0$, and hence $[0,\ -x^T]^T \in \mathcal{H}(D)$ so that $[0,\ x^T]^T \in -\mathcal{H}(D)$. $\square$

In the sequel of this section, we allow multiple quadratic constraints in the definition of D. In the next lemma, we impose a condition under which D must be bounded and, hence, relation (5) applies.

Lemma 3. *Let $q_i(x) = c_i + 2b_i^T x + x^T A_i x$, $i = 1, \ldots, m$. Assume that*

$$D = \{x \mid q_i(x) \geq 0,\ i = 1, \ldots, m\} \neq \varnothing.$$

Suppose furthermore that there exist $y_i \geq 0$, $i = 1, \ldots, m$, such that $\sum_{i=1}^m y_i A_i \prec 0$. In particular, this implies that D is a compact set. Then, we have

$$\mathcal{H}(D) = \left\{ \begin{bmatrix} t \\ x \end{bmatrix} \,\middle|\, t \geq 0,\ t^2 c_i + 2t b_i^T x + x^T A_i x \geq 0,\ i = 1, \ldots, m \right\}.$$

Proof. We first remark that $x \in D$ implies $\sum_{i=1}^m y_i q_i(x) \geq 0$. Because $\sum_{i=1}^m y_i q_i(x)$ is a strictly concave quadratic function, it follows that D is (contained in) a bounded set. Because $\sum_{i=1}^m y_i A_i \prec 0$ and $y_i \geq 0$ for all $i = 1, \ldots, m$, we also have the obvious implication

$$\min_{i=1,\ldots,m} x^T A_i x \geq 0 \quad \Rightarrow \quad x^T \left(\sum_{i=1}^m y_i A_i \right) x \geq 0 \quad \Rightarrow \quad x = 0. \tag{9}$$

Therefore, we have

$$\begin{aligned}
&\left\{ \begin{bmatrix} t \\ x \end{bmatrix} \,\middle|\, t \geq 0,\ t^2 c_i + 2t b_i^T x + x^T A_i x \geq 0,\ i = 1, \ldots, m \right\} \\
&\quad = \left\{ \begin{bmatrix} t \\ x \end{bmatrix} \,\middle|\, t > 0,\ t^2 q_i(x/t) \geq 0,\ i = 1, \ldots, m \right\} \cup \left\{ \begin{bmatrix} 0 \\ x \end{bmatrix} \,\middle|\, x^T A_i x \geq 0,\ i = 1, \ldots, m \right\} \\
&\quad = \left\{ \begin{bmatrix} t \\ x \end{bmatrix} \,\middle|\, t > 0,\ x/t \in D \right\} \cup \{0\} \\
&\quad = \mathcal{H}(D),
\end{aligned}$$

where the last two steps follow from (9) and (5), respectively. $\square$

Because an equality constraint can be represented by two inequalities, we arrive at the following corollary.

Corollary 2. *Let $q_i(x) = c_i + 2b_i^T x + x^T A_i x$, $i = 1, \ldots, m + l$. Assume that*

$$D = \{x \mid q_i(x) \geq 0,\ i = 1, \ldots, m \text{ and } q_j(x) = 0,\ j = m+1, \ldots, m+l\} \neq \varnothing.$$

Suppose furthermore that there exist $y_i \geq 0$, $i = 1, \ldots, m + l$, such that $\sum_{i=1}^{m+l} y_i A_i \prec 0$. Then, we have

$$\begin{aligned}
\mathcal{H}(D) = \Bigg\{ \begin{bmatrix} t \\ x \end{bmatrix} \,\Bigg|\, & t \geq 0,\ t^2 c_i + 2t b_i^T x + x^T A_i x \geq 0,\ i = 1, \ldots, m; \\
& t^2 c_j + 2t b_j^T x + x^T A_j x = 0,\ j = m+1, \ldots, m+l \Bigg\}.
\end{aligned}$$

The next lemma deals with a convex domain. In the presence of convexity, we no longer require D to be bounded. As a special case, it includes a domain defined by one concave and one linear inequality; this case will be studied in detail later.

LEMMA 4. *Let $q_i(x) = c_i + 2b_i^T x + x^T A_i x$, $i = 1, \ldots, m$, be concave functions. Suppose that*

$$D = \{x \mid q_i(x) \geq 0,\ i = 1, \ldots, m\} \neq \varnothing.$$

We have

$$\mathcal{H}(D) = \left\{ \begin{bmatrix} t \\ x \end{bmatrix} \middle| \, t \geq 0,\ t^2 c_i + 2tb_i^T x + x^T A_i x \geq 0,\ tc_i + 2b_i^T x \geq 0,\ i = 1, \ldots, m \right\}.$$

PROOF. For $[t,\ x^T]^T \in \mathcal{H}(D)$, we have a sequence $t_n > 0$, $x_n/t_n \in D$ with $(t_n, x_n) \to (t, x)$. The fact that $x_n/t_n \in D$ implies for all $i = 1, \ldots, m$ that

$$t_n > 0, \quad t_n^2 q_i(x_n/t_n) = t_n^2 c_i + 2t_n b_i^T x_n + x_n^T A_i x_n \geq 0 \tag{10}$$

and hence, using the concavity of $q_i(\cdot)$,

$$t_n c_i + 2b_i^T x_n \geq -x_n^T A_i x_n / t_n \geq 0. \tag{11}$$

By taking limits in the relations (10) and (11), we have

$$t \geq 0, \quad t^2 c_i + 2tb_i^T x + x^T A_i x \geq 0, \quad tc_i + 2b_i^T x \geq 0, \quad i = 1, \ldots, m. \tag{12}$$

Conversely, assume that (12) holds. If $t > 0$, then (12) implies that $t^2 q_i(x/t) \geq 0$ for all $i = 1, \ldots, m$, so that $x/t \in D$ and hence $[t,\ x^T]^T \in \mathcal{H}(D)$. Otherwise, i.e. if $t = 0$, then (12) implies that $x^T A_i x \geq 0$ and $b_i^T x \geq 0$ for all $i = 1, \ldots, m$. Because the A_is are negative semidefinite, it further follows that $A_i x = 0$ for all $i = 1, \ldots, m$. Therefore, we have for $\bar{x} \in D$ and $\epsilon > 0$, that

$$q_i(\bar{x} + x/\epsilon) = q_i(\bar{x}) + 2b_i^T x/\epsilon \geq 0 \quad \text{for all } i = 1, \ldots, m,$$

and hence $[\epsilon,\ x^T + \epsilon \bar{x}^T]^T \in \mathcal{H}(D)$. Letting $\epsilon \downarrow 0$, it follows that $[0,\ x^T]^T \in \mathcal{H}(D)$, as desired. □

Interestingly, $\mathcal{H}(D)$ in Lemma 4 admits a second-order cone representation (Nesterov and Nemirovsky 1994):

LEMMA 5. *Let $q(x) = c + 2b^T x + x^T A x$ be a concave function, and R a matrix such that $A = -R^T R$. Let r and n denote the number of rows and columns in R, respectively. The following three statements for $t \in \Re$ and $x \in \Re^n$,* (13), (14), *and* (15), *are equivalent*:

$$t \geq 0, \quad t^2 c + 2tb^T x + x^T A x \geq 0, \quad tc + 2b^T x \geq 0, \tag{13}$$

$$(ct + 2b^T x + t) \geq \sqrt{(ct + 2b^T x - t)^2 - 4x^T A x}, \tag{14}$$

$$\begin{bmatrix} c+1, & 2b^T \\ c-1, & 2b^T \\ 0, & 2R \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix} \in \mathrm{SOC}(r+2). \tag{15}$$

PROOF. Statements (14) and (15) are obviously equivalent. It remains to show that they are also equivalent with (13). Observe that in general, for any $\alpha, \beta \in \Re$, we have

$$\alpha + \beta \geq |\alpha - \beta| \iff \alpha \geq 0, \quad \beta \geq 0.$$

Therefore, we have in particular that

$$ct + 2b^T x + t \geq |ct + 2b^T x - t| \iff t \geq 0, \quad tc + 2b^T x \geq 0.$$

The correctness of the lemma is now easily verified. □

The advantage of having a second-order cone formulation of $\mathscr{H}(D)$, $D=\{x \mid q(x)\geq 0\}$ with $q(\cdot)$ concave, is that we immediately also get a second-order cone formulation of the dual cone, $\mathscr{H}(D)^*$. Namely, we have in general for a given $k\times n$ matrix B and a cone K that

$$(16)\qquad Bx\in K^* \iff y^T Bx\geq 0\,\forall y\in K \iff x\in\{B^T y\mid y\in K\}^*,$$

i.e.,

$$(17)\qquad \{x\mid Bx\in K^*\}=\{B^T y\mid y\in K\}^*;$$

when x is not a vector but a matrix, we may either use the above identity after vectorization, or interpret B^T as the adjoint of a linear operator B.

Corollary 3. *Let $q(x)=c+2b^T x+x^T Ax$ be a concave function with $D=\{x\mid q(x)\geq 0\}\neq\varnothing$. Then,*

$$\mathscr{H}(D)=\{x\mid Bx\in \mathrm{SOC}(2+\operatorname{rank}(A))\}=\{B^T y\mid y\in\mathrm{SOC}(2+\operatorname{rank}(A))\}^*,$$

where

$$B=\begin{bmatrix} c+1, & 2b^T\\ c-1, & 2b^T\\ 0, & 2R\end{bmatrix},$$

and R is a $\operatorname{rank}(A)\times n$ *matrix such that $A=-R^T R$.*

For the case that b is in the image of A, or equivalently $\sup\{q(x)\mid x\in\Re^n\}<\infty$, a more compact second-order cone formulation is possible:

Lemma 6. *Let $q(x)=c+2b^T x+x^T Ax$ be a concave function with $b=A\beta$. Let R be an $r\times n$ matrix such that $A=-R^T R$. Then,* $\max\{q(x)\mid x\in\Re^n\}=c-b^T\beta$. *If $c-b^T\beta\geq 0$, then*

$$t\geq 0,\quad t^2c+2tb^T x+x^T Ax\geq 0,\quad tc+2b^T x\geq 0$$

if and only if

$$(18)\qquad \begin{bmatrix}\sqrt{c-b^T\beta}, & 0\\ R\beta, & R\end{bmatrix}\begin{bmatrix} t\\ x\end{bmatrix}\in\mathrm{SOC}(r+1).$$

5. Dual matrix decompositions. This section addresses the problem of computing the Carathéodory decomposition of a dual matrix solution $Z\in\mathscr{F}\mathscr{C}_+(D)^*$ into rank-1 solutions $Z=\sum_{i=1}^{N+1} y_i y_i^T$, $y_i\in\mathscr{H}(D)$. Moreover, we will use the decomposition algorithms in this section to obtain LMI characterizations of cones of nonnegative quadratic functions over certain quadratically constrained regions.

5.1. One quadratic constraint. In this subsection, we are concerned with a domain given by an upper level set of a single quadratic function. We shall first discuss a relatively simple matrix decomposition problem. A derivation of the LMI characterization of $\mathscr{F}\mathscr{C}_+(D)$ follows thereafter.

As is well known, a matrix $X\in\mathscr{S}^{n\times n}$ is a positive semidefinite matrix of rank r if and only if there exist $p_i\in\Re^n$, $i=1,2,\ldots,r$, such that

$$X=\sum_{i=1}^{r} p_i p_i^T.$$

A somewhat unexpected result is the following equivalence statement.

PROPOSITION 3. *Let $X \in \mathcal{S}^{n\times n}$ be a positive semidefinite matrix of rank r. Let $G \in \mathcal{S}^{n\times n}$ be a given matrix. Then, $G \bullet X \geq 0$ if and only if there exist $p_i \in \Re^n$, $i = 1, 2, \ldots, r$, such that*

$$X = \sum_{i=1}^{r} p_i p_i^T \quad \text{and} \quad p_i^T G p_i \geq 0 \qquad \text{for all } i = 1, 2, \ldots, r.$$

Our proof for this proposition is constructive. The crux of the construction is highlighted in the following procedure, which is easily implementable. We remark that it is possible to prove the decomposition result as stated in Proposition 3 by other means. One alternative proof, e.g., is to apply Theorem 2.1 in Pataki (1998). Furthermore, we will show that it is, to some extent, dual to the so-called S-lemma (Yakubovich 1971).

PROCEDURE 1.
Input: $X, G \in \mathcal{S}^{n\times n}$ *such that* $0 \neq X \succeq 0$ *and* $G \bullet X \geq 0$.
Output: Vector $y \in \Re^n$ *with* $0 \leq y^T G y \leq G \bullet X$ *such that* $X - yy^T$ *is a positive semidefinite matrix of rank* $r - 1$ *where* $r = \operatorname{rank}(X)$.

Step 0. Compute $p_1, \ldots, p_r$ such that $X = \sum_{i=1}^r p_i p_i^T$.
Step 1. If $(p_1^T G p_1)(p_i^T G p_i) \geq 0$ for all $i = 2, 3, \ldots, r$ then return $y = p_1$. Otherwise, let j be such that $(p_1^T G p_1)(p_j^T G p_j) < 0$.
Step 2. Determine α such that $(p_1 + \alpha p_j)^T G (p_1 + \alpha p_j) = 0$. Return $y = (p_1 + \alpha p_j)/\sqrt{1+\alpha^2}$.

LEMMA 7. *Procedure* 1 *is correct.*

PROOF. If the procedure stops in Step 1 with $y = p_1$, then all the quantities $p_i^T G p_i$, $i = 1, \ldots, r$, have the same sign. Furthermore, the sum of these quantities is nonnegative, because

$$\sum_{i=1}^{r} p_i^T G p_i = G \bullet X \geq 0.$$

Therefore, $p_i^T G p_i \geq 0$ for all $i = 1, 2, \ldots, r$. Moreover, $X - yy^T = \sum_{i=2}^r p_i p_i^T$ so that indeed $X - yy^T$ is a positive semidefinite matrix of rank $r - 1$.

Otherwise (i.e., the procedure does not stop in Step 1), the quadratic equation in Step 2 of Procedure 1 always has 2 distinct roots, because $(p_1^T G p_1)(p_j^T G p_j) < 0$. The definitions of α and y in Step 2 imply that $0 = y^T G y \leq G \bullet X$. Moreover, by letting $u := (p_j - \alpha p_1)/\sqrt{1+\alpha^2}$, we have

$$X - yy^T = uu^T + \sum_{i \in \{2, 3, \ldots, r\} \setminus j} p_i p_i^T,$$

which has rank $r - 1$, establishing the correctness of Procedure 1. □

PROOF OF PROPOSITION 3. It is obvious that the statement holds true for a matrix X of rank 0. Assume now that such is true for any matrix X with $\operatorname{rank}(X) \in \{0, 1, \ldots, r\}$ for a certain $r \in \{0, 1, \ldots, n-1\}$. Consider $X \in \mathcal{S}_+^{n\times n}$ with $G \bullet X \geq 0$ and $\operatorname{rank}(X) = r + 1$. Applying Procedure 1, and using Lemma 7, we can find y_1 such that

$$\operatorname{rank}(X - y_1 y_1^T) = r, \quad X - y_1 y_1^T \succeq 0, \quad 0 \leq y_1^T G y_1 \leq G \bullet X.$$

By induction, we conclude that there exist $y_2, \ldots, y_{r+1}$ such that

$$X - y_1 y_1^T = \sum_{i=2}^{r+1} y_i y_i^T$$

where $y_i^T G y_i \geq 0$, $i = 2, \ldots, r+1$. □

Proposition 3 can be readily extended to a more specific form, as shown in the following corollary.

COROLLARY 4. *Let $X \in \mathcal{S}^{n\times n}$ be a positive semidefinite matrix of rank r. Let $G \in \mathcal{S}^{n\times n}$ be a given matrix, and $G \bullet X \geq 0$. Then, we can always find $p_i \in \Re^n$, $i = 1, 2, \ldots, r$, such that*

$$X = \sum_{i=1}^{r} p_i p_i^T \quad \text{and} \quad p_i^T G p_i = G \bullet X/r \quad \text{for } i = 1, 2, \ldots, r.$$

The key to note here is that if $p_i^T G p_i = G \bullet X/r$ are not satisfied for all $i = 1, \ldots, r$, then there will always exist two indices, say i and j, such that $p_i^T G p_i < G \bullet X/r$ and $p_j^T G p_j > G \bullet X/r$. Similar to Procedure 1, we can always find α, such that $(p_i + \alpha p_j)^T G(p_i + \alpha p_j) = G \bullet X/r$. Below we shall use the decomposition result in Proposition 3 to get explicit representations of some nonnegative quadratic cones. We will use the property that if K_1 and K_2 are two convex cones, then

$$K_1^* \cap K_2^* = (K_1 + K_2)^*, \tag{19}$$

where $K_1 + K_2 = \{x + y \mid x \in K_1,\ y \in K_2\}$ (see Corollary 16.4.2 in Rockafellar 1970). In fact, (19) is a special case of (17) with $K = K_1 \times K_2$ and $B = [I,\ \ I]$. Dualizing both sides of (19), we also have (using the bi-polar theorem)

$$(K_1^* \cap K_2^*)^* = \operatorname{cl}(K_1 + K_2). \tag{20}$$

THEOREM 1. *Let q: $\Re^n \to \Re$ be a quadratic function, and suppose that the upper level set $D = \{x \mid q(x) \geq 0\}$ is nonempty. Then,*

$$\operatorname{conv}\left\{ yy^T \,\middle|\, y \in \mathcal{H}(D) \right\} = \{ X \succeq 0 \mid \ \mathrm{M}(q(\cdot)) \bullet X \geq 0 \}. \tag{21}$$

The cone of quadratic functions that are nonnegative on D is therefore

$$\mathcal{FC}_+(D) = \{ X \succeq 0 \mid \ \mathrm{M}(q(\cdot)) \bullet X \geq 0 \}^* = \operatorname{cl}\{ Z \mid Z - t\ \mathrm{M}(q(\cdot)) \succeq 0,\ t \geq 0 \}. \tag{22}$$

PROOF. Using Proposition 3 and Lemma 2 respectively, we have

$$\begin{aligned} \{X \succeq 0 \mid \ \mathrm{M}(q(\cdot)) \bullet X \geq 0\} &= \operatorname{conv}\left\{ yy^T \,\middle|\, y^T\ \mathrm{M}(q(\cdot)) y \geq 0 \right\} \\ &= \operatorname{conv}\left\{ yy^T \,\middle|\, y \in \mathcal{H}(D) \cup (-\mathcal{H}(D)) \right\}, \end{aligned}$$

and obviously $\operatorname{conv}\{yy^T \mid y \in \mathcal{H}(D) \cup (-\mathcal{H}(D))\} = \operatorname{conv}\{yy^T \mid y \in \mathcal{H}(D)\}$. This establishes (21).

Using Corollary 1 and relation (21), we have

$$\mathcal{FC}_+(D) = \operatorname{conv}\{yy^T \mid y \in \mathcal{H}(D)\}^* = \{X \succeq 0 \mid \ \mathrm{M}(q(\cdot)) \bullet X \geq 0\}^*. \tag{23}$$

Applying (20), it further follows that

$$\begin{aligned} \mathcal{FC}_+(D) &= \ \operatorname{cl}\left(\mathcal{S}_+^{(1+n)\times(1+n)} + \{t\ \mathrm{M}(q(\cdot)) \mid t \geq 0\} \right) \\ &= \ \operatorname{cl}\{Z \mid Z - t\ \mathrm{M}(q(\cdot)) \succeq 0,\ t \geq 0\}. \quad \square \end{aligned}$$

We remark that in general, the set $\{Z \mid Z - t\ \mathrm{M}(q(\cdot)) \succeq 0,\ t \geq 0\}$ is not necessarily closed. Consider, for instance, the function q: $\Re \to \Re$ defined as $q(x) = -x^2$, for which

$D = \{x \mid q(x) \geq 0\} = \{0\}$. Clearly, the function $f(x) = 2x$ is nonnegative on D, but the 2×2 matrix

$$\mathrm{M}(f(\cdot)) - t\ \mathrm{M}(q(\cdot)) = \begin{bmatrix} 0, & 1 \\ 1, & 0 \end{bmatrix} - t \begin{bmatrix} 0, & 0 \\ 0, & -1 \end{bmatrix}$$

is not positive semidefinite for any t. However, for any $\epsilon > 0$ and $t \geq 1/\epsilon$, we have

$$\begin{bmatrix} \epsilon, & 1 \\ 1, & 0 \end{bmatrix} - t \begin{bmatrix} 0, & 0 \\ 0, & -1 \end{bmatrix} \succeq 0.$$

Letting $\epsilon \downarrow 0$, we see in this case that $\mathrm{M}(f(\cdot))$ is (merely) a limit point of

$$\{Z \mid Z - t\ \mathrm{M}(q(\cdot)) \succeq 0, t \geq 0\}.$$

As a corollary to Theorem 1, we arrive at the following well-known result from robust control, which is known as the S-procedure (Yakubovich 1971).

Corollary 5. *Let f: $\Re^n \to \Re$ and q: $\Re^n \to \Re$ be quadratic functions, and suppose that there exists $\bar{x} \in \Re^n$ such that $q(\bar{x}) > 0$. Let $D = \{x \mid q(x) \geq 0\}$. Then,*

$$\mathscr{F}\mathscr{C}_+(D) = \{Z \mid Z - t\ \mathrm{M}(q(\cdot)) \succeq 0,\ t \geq 0\}.$$

This means that $f(x) \geq 0$ for all $x \in D$ if and only if there exists $t \geq 0$ such that $f(x) - tq(x) \geq 0$ for all $x \in \Re^n$.

Proof. Let $y := [1,\ \bar{x}^T]^T$ and let

$$Z \in \mathrm{cl}\{Z \mid Z - t\ \mathrm{M}(q(\cdot)) \succeq 0,\ t \geq 0\}.$$

Then, there exist $Z_k \in \mathscr{S}^{(1+n)\times(1+n)}$ and $t_k \in \Re_+$ with $Z_k - t_k\ \mathrm{M}(q(\cdot)) \succeq 0$ and $Z_k \to Z$. We have

$$0 \leq y^T(Z_k - t_k\ \mathrm{M}(q(\cdot)))y = y^T Z_k y - t_k q(\bar{x}),$$

so that $0 \leq t_k \leq y^T Z_k y / q(\bar{x})$. It follows that $\{t_k\}$ is bounded, and hence it has a cluster point t such that $Z - t\ \mathrm{M}(q(\cdot)) \succeq 0$. This shows that

$$\{Z \mid Z - t\ \mathrm{M}(q(\cdot)) \succeq 0,\ t \geq 0\} \tag{24}$$

is closed.

By definition, $f(x) \geq 0$ for all $x \in D$ if and only if

$$\mathrm{M}(f(\cdot)) \in \mathscr{F}\mathscr{C}_+(D). \tag{25}$$

Using (7), we know that $f(x) - tq(x) \geq 0$ for all $x \in \Re^n$ if and only if

$$\mathrm{M}(f(\cdot)) - t\ \mathrm{M}(q(\cdot)) \in \mathscr{F}\mathscr{C}_+(\Re^n) = \mathscr{S}_+^{(1+n)\times(1+n)}. \tag{26}$$

Using Theorem 1 with (24), we have (25) if and only if (26) holds for some $t \geq 0$. □

The regularity condition that there exists $\bar{x} \in \Re^n$ such that $q(\bar{x}) > 0$ is equivalent to stating that $\mathrm{M}(q(\cdot))$ is not negative semidefinite. Namely, $q(x) \leq 0$ for all x if and only if $-q(\cdot)$ is nonnegative on the whole $\Re^n$, which holds if and only if $\mathrm{M}(q(\cdot)) \preceq 0$; see (7).

For the special case that $q(\cdot)$ is concave, the LMI representation of $\mathscr{F}\mathscr{C}_+(D)$ as stated in Theorem 1 can also be found in Fu et al. (1998) and Rendl and Wolkowicz (1997).

5.2. One quadratic equality constraint. The following proposition states a special case of Corollary 4.

PROPOSITION 4. *Let $X \in \mathcal{S}^{n\times n}$ be a positive semidefinite matrix of rank r. Let $G \in \mathcal{S}^{n\times n}$ be a given matrix. Then, $G \bullet X = 0$ if and only if there exist $p_i \in \Re^n$, $i = 1, 2, \ldots, r$, such that*

$$X = \sum_{i=1}^{r} p_i p_i^T \quad \textit{and} \quad p_i^T G p_i = 0 \quad \textit{for all } i = 1, 2, \ldots, r.$$

Similar to Theorem 1 we obtain from Proposition 4 and Corollary 2 the following result.

THEOREM 2. *Let q: $\Re^n \to \Re$ be a strictly concave quadratic function, and suppose that the level set $D = \{x \mid q(x) = 0\}$ is nonempty. Then,*

$$\operatorname{conv}\left\{yy^T \,\middle|\, y \in \mathcal{H}(D)\right\} = \left\{X \succeq 0 \,\middle|\, \mathrm{M}(q(\cdot)) \bullet X = 0\right\}$$

and

$$\mathcal{FC}_+(D) = \operatorname{cl}\left\{Z \mid Z - t\,\mathrm{M}(q(\cdot)) \succeq 0,\ t \in \Re\right\}.$$

COROLLARY 6. *Let f: $\Re^n \to \Re$ and q: $\Re^n \to \Re$ be quadratic functions, and suppose that $q(\cdot)$ is strictly concave and that there exists $\bar{x} \in \Re^n$ such that $q(\bar{x}) > 0$. Let $D = \{x \mid q(x) = 0\}$. Then, $f(x) \geq 0$ for all $x \in D$ if and only if there exists $t \in \Re$ such that $f(x) - tq(x) \geq 0$ for all $x \in \Re^n$.*

We remark that the existence of x such that $q(x) < 0$ follows from the strict concavity of $q(\cdot)$. The proof of Corollary 6 is analogous to the proof of Corollary 5.

Considering both Corollary 5 and Corollary 6, we remark that if a quadratic function $f(\cdot)$ is nonnegative on the level set $D = \{x \mid q(x) = 0\}$ of a strictly concave quadratic function $q(\cdot)$, then there cannot exist two solutions $x^{(1)}$ and $x^{(2)}$ such that $q(x^{(1)}) < 0$ and $q(x^{(2)}) > 0$, but $\max(f(x^{(1)}), f(x^{(2)})) < 0$.

5.3. One linear and one concave quadratic constraint. In this subsection, we will deal with a domain defined by one linear and one concave quadratic constraint.

Let $q(x) = c + 2b^T x + x^T Ax$ be a concave quadratic function with a nonempty upper level set $D := \{x \mid q(x) \geq 0\}$. Because of the concavity of $q(\cdot)$, D is convex and hence $\mathcal{H}(D)$ is a convex cone. Using Lemma 4, we have

$$\mathcal{H}(D) = \left\{\begin{bmatrix} t \\ x \end{bmatrix} \,\middle|\, t \geq 0,\ t^2 c + 2tb^T x + x^T Ax \geq 0,\ tc + 2b^T x \geq 0\right\}.$$

Due to the concavity of $q(\cdot)$, it holds that $x^T Ax \leq 0$ for all x and therefore

$$\begin{cases} t \geq 0,\ t^2 c + 2tb^T x + x^T Ax > 0 & \Rightarrow \quad t > 0,\ tc + 2b^T x > 0, \\ t > 0,\ t^2 c + 2tb^T x + x^T Ax \geq 0 & \Rightarrow \quad tc + 2b^T x \geq 0. \end{cases} \tag{27}$$

Suppose that $a \in \Re^{1+n}$ and $X \in \mathcal{S}_+^{(1+n)\times(1+n)}$ are such that $Xa \neq 0$. Let U be a matrix of full column rank such that

$$X = UU^T.$$

Then, we have $Xa = U(U^T a)$ so that

$$X - \frac{1}{a^T Xa}(Xa)(Xa)^T = U\left(I - \frac{1}{\|U^T a\|^2}(U^T a)(U^T a)^T\right)U^T. \tag{28}$$

It is clear that the right-hand side in the above equation is a positive semidefinite matrix of rank $r-1$, where $r = \operatorname{rank}(X)$. This fact is used in Lemma 8.

PROCEDURE 2.

Input: $X \in \mathcal{S}^{(1+n)\times(1+n)}$, *a concave quadratic function* $q: \Re^n \to \Re$ *and a vector* $a \in \Re^{1+n}$ *such that* $X \succeq 0$, $\mathrm{M}(q(\cdot)) \bullet X \geq 0$, *and* $0 \neq Xa \in \mathcal{H}(D)$, *where* $D := \{x \mid q(x) \geq 0\}$.

Output: *One of the two possibilities*:

- *Vector* $y \in \mathcal{H}(D)$ *with* $0 \leq y^T\ \mathrm{M}(q(\cdot))y \leq \mathrm{M}(q(\cdot)) \bullet X$ *and* $a^T y \geq 0$ *such that* $X^{\text{new}} := X - yy^T$ *is a positive semidefinite matrix of rank* $r-1$ *where* $r = \operatorname{rank}(X)$, *and* $X^{\text{new}} a \in \mathcal{H}(D)$.
- *Vector* $0 \neq y \in \mathcal{H}(D)$ *with* $0 \leq y^T\ \mathrm{M}(q(\cdot))y \leq \mathrm{M}(q(\cdot)) \bullet X$ *and* $a^T y \geq 0$ *such that* $X^{\text{new}} := X - yy^T$ *is a positive semidefinite matrix, and* $X^{\text{new}} a \neq 0$ *is on the boundary of* $\mathcal{H}(D)$, *i.e.*, $a^T X^{\text{new}}\ \mathrm{M}(q(\cdot)) X^{\text{new}} a = 0$.

Step 0. Let $p_1 := Xa/\sqrt{a^T X a}$ and compute $p_2, \ldots, p_r$ such that $X - p_1 p_1^T = \sum_{i=2}^r p_i p_i^T$ and that the first entry of p_i is nonnegative, $i = 2, \ldots, r$.

Step 1. If $p_1^T\ \mathrm{M}(q(\cdot))p_1 \leq \mathrm{M}(q(\cdot)) \bullet X$, then return $y = p_1$. Otherwise, let $j \in \{2, 3, \ldots, r\}$ be such that $p_j^T\ \mathrm{M}(q(\cdot))p_j < 0$.

Step 2. Determine $\alpha > 0$ such that $(p_1 + \alpha p_j)^T\ \mathrm{M}(q(\cdot))(p_1 + \alpha p_j) = 0$. Let

$$v = (p_1 + \alpha p_j)/\sqrt{1+\alpha^2} \quad \text{and} \quad w(t) = Xa - t(a^T v)v.$$

Define γ_0 and γ_1 to be such that $w(t)^T\ \mathrm{M}(q(\cdot))w(t) = \gamma_0 - \gamma_1 t$ for all $t \in \Re$.

Step 3. If $\gamma_1 > \gamma_0$, then let $y = \sqrt{\gamma_0/\gamma_1}\, v$, else let $y = v$.

LEMMA 8. *Procedure* 2 *is correct.*

PROOF. By definition of $\mathcal{H}(D)$, $Xa \in \mathcal{H}(D)$ implies $(Xa)^T\ \mathrm{M}(q(\cdot))Xa \geq 0$. Therefore, if Procedure 2 stops with $y = p_1 := Xa/\sqrt{a^T X a}$ in Step 1, then $y^T\ \mathrm{M}(q(\cdot))y = (Xa)^T\ \mathrm{M}(q(\cdot))Xa/(a^T Xa) \geq 0$, and $\operatorname{rank}(X - yy^T) = r - 1$ as stipulated by (28). Moreover, $a^T y = \sqrt{a^T X a} \geq 0$ so that $X^{\text{new}} a = Xa - (a^T y)y = Xa - Xa = 0 \in \mathcal{H}(D)$. Therefore, the procedure terminates correctly in Step 1.

Suppose now that the procedure does not stop at Step 1, i.e.

$$p_1^T\ \mathrm{M}(q(\cdot))p_1 > \mathrm{M}(q(\cdot)) \bullet X \geq 0. \tag{29}$$

Using (27), it follows that the first entry of p_1 is (strictly) positive. Furthermore, because

$$\sum_{j=2}^r p_j^T\ \mathrm{M}(q(\cdot))p_j = \mathrm{M}(q(\cdot)) \bullet X - p_1^T\ \mathrm{M}(q(\cdot))p_1 < 0,$$

it also follows that there is indeed a $j \in \{2, 3, \ldots, r\}$ such that $p_j^T\ \mathrm{M}(q(\cdot))p_j < 0$.

The quadratic equation in Step 2 of Procedure 2 always has one positive and one negative root, due to $p_1^T\ \mathrm{M}(q(\cdot))p_1 > 0$ and $p_j^T\ \mathrm{M}(q(\cdot))p_j < 0$. The procedure defines α to be the positive root. Because the first entry in p_j was made nonnegative in Step 0, it follows that $p_1 + \alpha p_j \in \Re_{++} \times \Re^n$. This further means that the first entry in $v := (p_1 + \alpha p_j)/\sqrt{1+\alpha^2}$ is positive. Moreover, $v^T\ \mathrm{M}(q(\cdot))v = 0$ due to the definition of α. As can be seen from (27), these two properties of v imply that $0 \neq v \in \mathcal{H}(D)$. This proves that $0 \neq y \in \mathcal{H}(D)$ after termination in Step 3.

Using the definition of p_1 and $p_2, p_3, \ldots, p_r$, we have

$$\left(\sum_{i=2}^r p_i p_i^T\right) a = (X - p_1 p_1^T)a = Xa - Xa = 0.$$

This implies that $\sum_{i=2}^{r}(p_i^T a)^2 = 0$ and hence $p_i^T a = 0$ for all $i = 2, \ldots, r$. We further obtain that

$$a^T v = \frac{a^T(p_1 + \alpha p_j)}{\sqrt{1+\alpha^2}} = \frac{a^T p_1}{\sqrt{1+\alpha^2}} = \sqrt{\frac{a^T X a}{1+\alpha^2}} > 0,$$

and hence $a^T y \geq 0$ after termination in Step 3.

The scalars γ_0 and γ_1 in Step 2 are well defined, because $v^T\ \mathrm{M}(q(\cdot))v = 0$ due to the definition of α. In fact, it is easily verified that $\gamma_0 = a^T X\ \mathrm{M}(q(\cdot))Xa > 0$ and $\gamma_1 = 2(a^T v)(a^T X\ \mathrm{M}(q(\cdot))v)$. To simplify notation, we define

$$\tau := \begin{cases} 1, & \text{if } \gamma_1 \leq \gamma_0, \\ \gamma_0/\gamma_1, & \text{if } \gamma_1 > \gamma_0\,(> 0), \end{cases}$$

so that $y = \sqrt{\tau}v$. By definition of γ_0, γ_1, and τ, we have

$$w(t)^T\ \mathrm{M}(q(\cdot))w(t) = \gamma_0 - t\gamma_1 > 0 \quad \text{for } 0 \leq t < \tau. \tag{30}$$

Using (27), this implies by a continuity argument that $w(\tau) \in \mathscr{H}(D)$. However,

$$w(\tau) = Xa - \tau(a^T v)v = Xa - (a^T y)y = X^{\mathrm{new}}a,$$

so that $X^{\mathrm{new}}a \in \mathscr{H}(D)$ as desired. Furthermore, we have

$$\tau < 1 \quad \Rightarrow \quad 0 = w(\tau)^T\ \mathrm{M}(q(\cdot))w(\tau) = a^T X^{\mathrm{new}}\ \mathrm{M}(q(\cdot))X^{\mathrm{new}}a,$$

which means that $X^{\mathrm{new}}a$ is on the boundary of $\mathscr{H}(D)$ if $\tau < 1$. Furthermore, $X^{\mathrm{new}}a = w(\tau) \neq 0$ because $w(t) \neq 0$ for any

It remains to verify that if $\tau = 1$, then $\operatorname{rank}(X - yy^T) = r - 1$, where $r = \operatorname{rank}(X)$. We now introduce $u = (p_j - \alpha p_1)/\sqrt{1+\alpha^2}$, for which we have the obvious relation

$$uu^T + vv^T = p_1 p_1^T + p_j p_j^T.$$

Because $\tau = 1$ implies $y = v$, we therefore get

$$X - yy^T = X - vv^T = uu^T + \sum_{i \in \{2,3,\ldots,r\} \setminus j} p_i p_i^T.$$

It follows that $\operatorname{rank}(X - yy^T) = r - 1$. $\square$

We observe from Lemma 8 that if Procedure 2 does *not* reduce the rank of X, then the vector $X^{\mathrm{new}}a$ is nonzero and on the boundary of $\mathscr{H}(D)$. However, if we apply the procedure to X^{new} we find that $0 \neq p_1^{\mathrm{new}} = X^{\mathrm{new}}a/\sqrt{a^T X^{\mathrm{new}} a}$ and $(p_1^{\mathrm{new}})^T\ \mathrm{M}(q(\cdot))p_1^{\mathrm{new}} = 0$. Therefore, the procedure exits at Step 1 to produce $X^{\mathrm{final}} := X^{\mathrm{new}} - p_1^{\mathrm{new}}(p_1^{\mathrm{new}})^T$ with $X^{\mathrm{final}}a = 0$. We will decompose X^{final} using Procedure 1. Based on this scheme, we arrive at the matrix decomposition result as stated in Proposition 5 below.

PROPOSITION 5. *Let $q: \Re^n \to \Re$ be a concave quadratic function, $D = \{x \mid q(x) \geq 0\} \neq \varnothing$, $X \in \mathscr{S}_+^{(1+n)\times(1+n)}$, and $\mathrm{M}(q(\cdot)) \bullet X \geq 0$, and a vector $a \in \Re^{1+n}$ be such that $Xa \in \mathscr{H}(D)$. Then, there exist y_i, $i = 1, \ldots, k$, for some $k \in \{r, r+1\}$ with $r = \operatorname{rank}(X)$, such that*

$$X = \sum_{i=1}^{k} y_i y_i^T$$

and $y_i \in \mathscr{H}(D)$, and $a^T y_i \geq 0$, $i = 1, \ldots, k$.

Proof. We distinguish three cases.

Case 1. If $Xa = 0$, then we invoke Procedure 1 to obtain $X = \sum_{i=1}^{r} y_i y_i^T$ with $y_i \in \mathscr{H}(D)$ for all $i = 1, 2, \ldots, r$ (see Proposition 3). Moreover, because $0 = a^T X a = \sum_{i=1}^{r} (a^T y_i)^2$, it follows that $a^T y_i = 0$ for $i = 1, 2, \ldots, r$. This shows that if $Xa = 0$, then the proposition holds with $k = r$.

Case 2. Now, consider the case that $Xa \neq 0$ and applying Procedure 2 once on X does *not* reduce the rank. Apply Procedure 2 to obtain $y_1 \in \mathscr{H}(D)$ with $a^T y_1 \geq 0$ such that

$$X^{\text{new}} := X - y_1 y_1^T \succeq 0, \quad X^{\text{new}} a \in \mathscr{H}(D), \quad \mathrm{M}(q(\cdot)) \bullet X^{\text{new}} \geq 0.$$

If $\operatorname{rank}(X^{\text{new}}) = \operatorname{rank}(X) = r$, then $a^T X^{\text{new}}\, \mathrm{M}(q(\cdot)) X^{\text{new}} a = 0$ and we can apply Procedure 2 on X^{new} to obtain $y_2 \in \mathscr{H}(D)$ with $a^T y_2 \geq 0$ such that

$$X^{\text{final}} := X - y_1 y_1^T - y_2 y_2^T \succeq 0, \quad X^{\text{final}} a = 0, \quad \mathrm{M}(q(\cdot)) \bullet X^{\text{final}} \geq 0,$$

and $\operatorname{rank}(X^{\text{final}}) = r - 1$. Because Case 1 applies to X^{final}, we know 7 that $X^{\text{final}} = \sum_{i=3}^{r+1} y_i y_i^T$, with $y_i \in \mathscr{H}(D)$ and $a^T y_i \geq 0$, $i = 3, 4, \ldots, r+1$. Hence, the proposition also holds true for Case 2 with $k = r + 1$.

Case 3. The remaining case is that $Xa \neq 0$ and applying Procedure 2 once on X reduces the rank. Because the rank is always nonnegative, we can reduce this case to either Case 1 or Case 2 by a recursive argument. We can now prove the proposition by induction on $\operatorname{rank}(X)$. Namely, suppose now that the proposition holds true for any matrix X with $\operatorname{rank}(X) \in \{0, 1, \ldots, r\}$ for a certain $r \in \{0, 1, \ldots, n\}$. Consider $X \in \mathscr{S}_+^{(n+1)\times(n+1)}$ with $\operatorname{rank}(X) = r + 1$, for which Procedure 2 yields a vector $y_1 \in \mathscr{H}(D)$ with $a^T y_1 \geq 0$ such that

$$X^{\text{new}} := X - y_1 y_1^T \succeq 0, \quad X^{\text{new}} a \in \mathscr{H}(D), \quad \mathrm{M}(q(\cdot)) \bullet X^{\text{new}} \geq 0,$$

and $\operatorname{rank}(X^{\text{new}}) = \operatorname{rank}(X) - 1 = r$. By induction, we conclude that there exist $y_2, \ldots, y_{k+1}$ for some $k \in \{r, r+1\}$ such that

$$X - y_1 y_1^T = \sum_{i=2}^{k+1} y_i y_i^T,$$

where $y_i \in \mathscr{H}(D)$ and $a^T y_i \geq 0$ for all $i = 2, \ldots, k+1$. □

Using similar reasoning as before, the above decomposition result implies an LMI characterization of $\mathscr{F}\mathscr{C}_+(D)$.

Theorem 3. *Let $q\colon \Re^n \to \Re$ be a concave quadratic function, $a \in \Re^{n+1}$. Let*

$$D := \{x \mid q(x) \geq 0\}, \quad L := \{x \mid [1, \quad x^T]\, a \geq 0\}.$$

Suppose $D \cap L \neq \varnothing$. Then,

$$\operatorname{conv}\{yy^T \mid y \in \mathscr{H}(D \cap L)\} = \{X \in \mathscr{S}_+^{(1+n)\times(1+n)} \mid \mathrm{M}(q(\cdot)) \bullet X \geq 0,\ Xa \in \mathscr{H}(D)\}.$$

Consequently, the cone of all quadratic functions that are nonnegative on $D \cap L$ is

$$\begin{aligned}\mathscr{F}\mathscr{C}_+(D \cap L) &= \left\{X \in \mathscr{S}_+^{(1+n)\times(1+n)} \mid \mathrm{M}(q(\cdot)) \bullet X \geq 0,\ Xa \in \mathscr{H}(D)\right\}^* \\ &= \operatorname{cl}\{Z \mid Z - (t\, \mathrm{M}(q(\cdot)) + a\psi^T + \psi a^T) \succeq 0,\ t \geq 0, \psi \in \mathscr{H}(D)^*\}.\end{aligned}$$

Proof. By Lemma 4, we know that

$$\mathscr{H}(D \cap L) = \{y \in \mathscr{H}(D) \mid a^T y \geq 0\},$$

and so

$$\operatorname{conv}\{yy^T \mid y \in \mathscr{H}(D \cap L)\} = \operatorname{conv}\{yy^T \mid y \in \mathscr{H}(D),\ a^T y \geq 0\}.$$

Suppose that X is a matrix in the above set, i.e. $X = \sum_{i=1}^{k} y_i y_i^T \succeq 0$ with $y_i \in \mathcal{H}(D)$ and $a^T y \geq 0$ for $i = 1, 2, \ldots, k$. Because $y_i \in \mathcal{H}(D)$, we certainly have $y_i^T \ \mathrm{M}(q(\cdot)) y_i \geq 0$ and consequently

$$M(q(\cdot)) \bullet X = \sum_{i=1}^{k} y_i^T \ \mathrm{M}(q(\cdot)) y_i \geq 0.$$

Moreover, $a^T y_i \geq 0$ and $y_i \in \mathcal{H}(D)$ for all $i = 1, 2, \ldots, k$, and

$$Xa = \sum_{i=1}^{k} (a^T y_i) y_i.$$

In other words, Xa is a nonnegative combination of vectors in the cone $\mathcal{H}(D)$, which implies that $Xa \in \mathcal{H}(D)$.

Conversely, for $X \succeq 0$ with $M(q(\cdot)) \bullet X \geq 0$ and $Xa \in \mathcal{H}(D)$, we know from Proposition 5 that $X = \sum_{i=1}^{k} y_i y_i^T$ with $y_i \in \mathcal{H}(D)$ and $a^T y \geq 0$ for $i = 1, 2, \ldots, k$. We conclude that

$$\text{(31)} \qquad \operatorname{conv}\left\{ yy^T \,\middle|\, y \in \mathcal{H}(D \cap L) \right\} = \left\{ X \in \mathcal{S}_+^{(1+n)\times(1+n)} \,\middle|\, \mathrm{M}(q(\cdot)) \bullet X \geq 0,\ Xa \in \mathcal{H}(D) \right\}.$$

Using Corollary 1 and (31), we have

$$\begin{aligned} \mathcal{F}\mathcal{C}_+(D \cap L) &= \operatorname{conv}\{ yy^T \mid y \in \mathcal{H}(D \cap L) \}^* \\ &= \left\{ X \in \mathcal{S}_+^{(1+n)\times(1+n)} \,\middle|\, \mathrm{M}(q(\cdot)) \bullet X \geq 0,\ Xa \in \mathcal{H}(D) \right\}^*. \end{aligned} \tag{32}$$

We remark from (17) that

$$\{ ay^T + ya^T \mid y \in \mathcal{K} \}^* = \left\{ X \in \mathcal{S}^{n\times n} \,\middle|\, Xa \in \mathcal{K}^* \right\}, \tag{33}$$

where the dual is taken in the Euclidean space $\mathcal{S}^{n\times n}$.

Applying (20) and (32)–(33), it follows that

$$\begin{aligned} \mathcal{F}\mathcal{C}_+(D \cap L) &= \operatorname{cl}\left(\mathcal{S}_+^{(1+n)\times(1+n)} + \{ t \ \mathrm{M}(q(\cdot)) \mid t \geq 0 \} + \{ a\psi + \psi a^T \mid \psi \in \mathcal{H}(D)^* \} \right) \\ &= \operatorname{cl}\left\{ Z \mid Z - (t \ \mathrm{M}(q(\cdot)) + a\psi^T + \psi a^T) \succeq 0,\ t \geq 0, \psi \in \mathcal{H}(D)^* \right\}. \quad \square \end{aligned}$$

We recall from Corollary 3 that

$$\mathcal{H}(D) = \{ x \mid Bx \in \mathrm{SOC}(2 + \operatorname{rank}(A)) \} = \{ B^T y \mid y \in \mathrm{SOC}(2 + \operatorname{rank}(A)) \}^*,$$

for a certain matrix B depending on A, b, c. Therefore, Theorem 1 characterizes $\mathcal{F}\mathcal{C}_+(D \cap L)$ and its dual in terms of semidefinite and second-order cone constraints. As a corollary to Theorem 1, we arrive at the following result.

COROLLARY 7. *Let $f\colon \Re^n \to \Re$ and $q\colon \Re^n \to \Re$ be quadratic functions, and $a \in \Re^{n+1}$. Suppose $q(\cdot)$ is concave and that there exists $\bar{x} \in \Re^n$ such that $q(\bar{x}) > 0$ and $[1,\ \ \bar{x}^T]a > 0$. Let*

$$D := \{ x \mid q(x) \geq 0 \}, \qquad L := \left\{ x \mid \begin{bmatrix} 1, & x^T \end{bmatrix} a \geq 0 \right\}.$$

Then,

$$\mathcal{F}\mathcal{C}_+(D \cap L) = \left\{ Z \mid Z - (t \ \mathrm{M}(q(\cdot)) + a\psi^T + \psi a^T) \succeq 0,\ t \geq 0, \psi \in \mathcal{H}(D)^* \right\}.$$

This means that $f(x) \geq 0$ for all $x \in D \cap L$, if and only if there exists $t \geq 0$ and $\psi \in \mathcal{H}(D)^*$ such that

$$f(x) - tq(x) - 2\left(\begin{bmatrix} 1, & x^T \end{bmatrix} a \right) \left(\begin{bmatrix} 1, & x^T \end{bmatrix} \psi \right) \geq 0 \qquad \textit{for all } x \in \Re^n.$$

PROOF. Let

$$Z \in \operatorname{cl}\left\{Z \mid Z - \left(t\ \mathrm{M}(q(\cdot)) + a\psi^T + \psi a^T\right) \succeq 0,\ t \geq 0, \psi \in \mathcal{H}(D)^*\right\}.$$

Then, there exist $Z_k \in \mathcal{S}^{(1+n)\times(1+n)}$, $t_k \in \Re_+$, and $\psi_k \in \mathcal{H}(D)^*$ such that

$$Z_k - (t_k\ \mathrm{M}(q(\cdot)) + a\psi_k^T + \psi_k a^T) \succeq 0, \qquad Z_k \to Z. \tag{34}$$

Let $y := [1,\ \bar{x}^T]^T$. Clearly, $a^T y > 0$ and $y^T\ \mathrm{M}(q(\cdot))y = q(\bar{x}) > 0$. Because $q(\bar{x}) > 0$, it follows that y is in the interior of $\mathcal{H}(D)$, and hence

$$\psi_k^T y > 0 \ \text{ for all } 0 \neq \psi_k \in \mathcal{H}(D)^*. \tag{35}$$

Due to (34), we have

$$0 \leq y^T\left(Z_k - t_k\ \mathrm{M}(q(\cdot)) - 2a\psi_k^T\right)y = y^T Z_k y - t_k q(\bar{x}) - 2(a^T y)(\psi_k^T y).$$

Now, using the fact that $q(\bar{x}) > 0$, $a^T y > 0$, and $\psi_k^T y \geq 0$, we obtain that

$$0 \leq t_k \leq y^T Z_k y / q(\bar{x}), \quad 0 \leq \psi_k^T y \leq y^T Z_k y/(2a^T y),$$

which shows that t_k and $\psi_k^T y$ are bounded. Furthermore, y is in the interior of the (solid) cone $\mathcal{H}(D)$, so that the facts that $\psi_k^T y$ is bounded and $\psi_k \in \mathcal{H}(D)^*$ implies that $\|\psi_k\|$ is bounded. Therefore, the sequences t_k and ψ_k have cluster points t and ψ respectively, and

$$Z - (t\ \mathrm{M}(q(\cdot)) + a\psi^T + \psi a^T) \succeq 0.$$

It follows that

$$\left\{Z \mid Z - (t\ \mathrm{M}(q(\cdot)) + a\psi^T + \psi a^T) \succeq 0,\ t \geq 0, \psi \in \mathcal{H}(D)^*\right\} \quad \text{is closed.} \tag{36}$$

The corollary now follows by the same argument as in the proof of Corollary 5. □

We remark that, using the problem formulation (MD) in §3, minimizing a quadratic function $f(\cdot)$ over the set $D \cap L$ can now be equivalently written as

$$\begin{array}{ll} \text{minimize} & \mathrm{M}(f(\cdot)) \bullet X \\ \text{subject to} & \mathrm{M}(q(\cdot)) \bullet X \geq 0, \\ & Xa \in \mathcal{H}(D), \\ & x_{11} = 1, \\ & X \succeq 0. \end{array}$$

This formulation, which is a *semidefinite programming* problem with the same optimal value as the original problem, is different from a straightforward semidefinite relaxation problem

$$\begin{array}{ll} \text{minimize} & \mathrm{M}(f(\cdot)) \bullet X \\ \text{subject to} & \mathrm{M}(q(\cdot)) \bullet X \geq 0, \\ & e_1^T X a \geq 0, \\ & x_{11} = 1, \\ & X \succeq 0. \end{array}$$

Notice that $e_1 a^T + a e_1^T$ is the matrix representation of the linear inequality, so that the above relaxation corresponds to applying the S-procedure with two quadratic constraints. This relaxation may admit a gap with the original problem. For instance, if $q(x) = 1 - x^2$,

$a=[0,\ 1]^T$, and $f(x)=1+x-x^2$, then the optimal solutions are $x=1$ or $x=0$ with value $f(x)=1$. However, the optimal solution to the straightforward semidefinite relaxation is $X=I$ with value $\mathrm{M}(f(\cdot))\bullet I=0$. Indeed, $Xa=e_2\notin \mathrm{SOC}(2)$ so that X cannot be decomposed as a convex combination of feasible rank-1 solutions.

6. Conclusion. The results claimed in Theorems 1, 2, and 3 are quite powerful. They characterize, using LMIs, all the quadratic functions that are nonnegative over the respectively specified domains. If we decompose the optimal solution using Procedures 1 and 2, we find that all the components y_i yield optimal solutions and directions to the (nonconvex) quadratic optimization problem. To the best of our knowledge, such decomposition procedures have not been proposed before.

In trust region methods for nonlinear programming, one often needs to solve problems of type (P) in §3, where D is a unit ball. The problem is known to be solvable in polynomial time; for detailed discussions, see Ye (1997). Our result extends the polynomial solvability property to a nonconvex quadratic constraint (inequality or equality) and a nonconvex quadratic objective. Another case that we can handle is a nonconvex objective with a concave quadratic inequality constraint and an additional linear restriction. The complexity status of the problem to minimize a nonconvex quadratic function over the intersection of two general ellipsoids is still an open problem in the study of trust region methods. However, our last application solves this problem for the special case where the two ellipsoids, or more generally, level sets of two concave quadratic functions, have the same geometric structure (but may still be of very different sizes). Specifically, consider

$$\begin{array}{ll} \text{minimize} & q_0(x) \\ \text{subject to} & q_1(x)=x^TQx-2b_1^Tx+c_1\geq 0, \\ & q_2(x)=x^TQx-2b_2^Tx+c_2\geq 0. \end{array}$$

The key to note is that the feasible set of the above problem can be viewed as the union of two sets

$$\{x \mid x^TQx-2b_1^Tx+c_1\geq 0\}\cap\{x\mid 2(b_2-b_1)^Tx+c_2-c_1\geq 0\}$$

and

$$\{x \mid x^TQx-2b_2^Tx+c_2\geq 0\}\cap\{x\mid 2(b_1-b_2)^Tx+c_1-c_2\geq 0\}.$$

Minimizing an indefinite quadratic function $q_0(x)$ over each of these sets individually can be solved via an SDP formulation as shown in this paper. Hence, applying the method twice solves the whole problem. (We thank Yinyu Ye for our discussions leading to this observation.)

Acknowledgments. Part of this research was performed by J. F. Sturm at the Department of Quantitative Economics, Maastricht University, Maastricht, The Netherlands, and part was performed when this author was a visiting researcher at the Chinese University of Hong Kong. S. Zhang was supported by Hong Kong RGC Earmarked Grant CUHK4181/00E.

References

Barvinok, A. I. 1995. Problems of distance geometry and convex properties of quadratic maps. *Discrete and Comput. Geometry* **13** 189–202.

Ben-Tal, A., A. S. Nemirovsky. 1998. Robust convex optimization. *Math. Oper. Res.* **23**(4) 769–805.

Carathéodory, C. 1907. Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen. *Mathematische Annalen* **64** 95–115.

de Klerk, E., D. V. Pasechnik. 2000. Approximating the stability number of a graph via copositive programming. Technical report, TU Delft, Faculty of Information Technology and Systems, Delft, The Netherlands.

Dines. L. L. 1941. On the mapping of quadratic forms. *Bull. Amer. Math. Soc.* **47** 494–498.

El Ghaoui, L., H. Lebret. 1997. Robust solutions to least-squares problems with uncertain data. *SIAM J. Matrix Anal. and Appl.* **18**(4) 1035–1064.

Fradkov, A. L., V. A. Yakubovich. 1979. The S-procedure and duality relations in nonconvex problems of quadratic programming. Leningrad, Russia, Vestnik Leningrad University, Vol. 6, 101–109. English translation; original Russian publication in Vestnik Leningradskogo Universiteta, Seriya Matematika, No. 1, 1973.

Fu, M., Z-Q. Luo, Y. Ye. 1998. Approximation algorithms for quadratic programming. *J. Combin. Optim.* **2**(1) 29–50.

Hiriart-Urruty, J-B., M. Torki. 2001. Pemanently going back and forth between the "quadratic world" and the "convexity world" in optimization. Technical report, Université Paul Sabatier, Toulouse Cedex, France.

Lasserre, J. B. 2001. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* **11**(3) 796–817.

Luo, Z-Q., J. F. Sturm. (2000). Error bounds for quadratic systems. J. B. G. Frenk, C. Roos, T. Terlaky, S. Zhang, eds. *High Performance Optimization*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 383–404.

Murty, K. G., S. N. Kabadi. 1987. Some NP-complete problems in quadratic and nonlinear programming. *Math. Programming* **39** 117–129.

Nesterov, Y. 2000. Squared functional systems and optimization problems. J. B. G. Frenk, C. Roos, T. Terlaky, S. Zhang, eds. *High Performance Optimization*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 405–440.

———, A. Nemirovsky. 1994. *Interior Point Polynomial Methods in Convex Programming*, Vol. 13. *Studies in Applied Mathematics*. SIAM, Philadelphia, PA.

Parrilo, P. A. 2000. On a decomposition of multivariable forms via LMI methods. Technical report, California Institute of Technology, Pasadena, CA.

Pataki, G. 1998. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res.* **23**(2) 339–358.

Polyak, B. T. 2001. Convexity of nonlinear image of a small ball with applications to optimization. *Set-Valued Analysis*. Forthcoming.

Quist, A. J., E. de Klerk, C. Roos, T Terlaky. 1998. Copositive relaxation for general quadratic programming. *Optim. Methods and Software* **9** 185–208. (Special Issue Celebrating the 60th Birthday of Professor Naum Shor.)

Rendl, F., H. Wolkowicz. 1997. A semidefinite framework for trust region subproblems with applications to large scale minimization. *Math. Programming* **77**(2) 273–299.

Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press, Princeton, NJ.

Sturm, J. F. 2000. Theory and algorithms for semidefinite programming. J. B. G. Frenk, C. Roos, T. Terlaky, S. Zhang, eds. *High Performance Optimization*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 3–196.

Vavavis, S. A. 1991. *Nonlinear Optimization: Complexity Issues*. Oxford Science, New York.

Wolkowicz, H., R. Saigal, L. Vandenberghe, eds. 2001. *Handbook on Semidefinite Programming*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Yakubovich, V. A. 1977. S-procedure in nonlinear control theory. Vestnik Leningrad University, Vol. 4, No. 1, 73–93. English translation; original Russian publication in Vestnik Leningradskogo Universiteta, Seriya Matematika, Leningrad, Russia, 1971, 62–77.

Ye, Y. 1997. *Interior Point Algorithms: Theory and Analysis*. John Wiley & Sons, New York.

J. F. Sturm: Department of Econometrics, Tilburg University, The Netherlands; e-mail: j.f.sturm@uvt.nl

S. Zhang: Department of System Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong; e-mail: zhang@se.cuhk.edu.hk