

QUADRATIC MATRIX PROGRAMMING*

AMIR BECK†

Abstract. We introduce and study a special class of nonconvex quadratic problems in which the objective and constraint functions have the form $f(\mathbf{X}) = \text{Tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) + 2\text{Tr}(\mathbf{B}^T \mathbf{X}) + c$, $\mathbf{X} \in \mathbb{R}^{n \times r}$. The latter formulation is termed *quadratic matrix programming* (QMP) of order r . We construct a specially devised semidefinite relaxation (SDR) and dual for the QMP problem and show that under some mild conditions strong duality holds for QMP problems with at most r constraints. Using a result on the equivalence of two characterizations of the nonnegativity property of quadratic functions of the above form, we are able to compare the constructed SDR and dual problems to other known SDRs and dual formulations of the problem. An application to robust least squares problems is discussed.

Key words. quadratic matrix programming, semidefinite relaxation, strong duality, nonconvex quadratic problems

AMS subject classifications. 90C20, 90C26, 90C46

DOI. 10.1137/05064816X

1. Introduction. This work is concerned with nonconvex quadratic optimization problems of the form

$$(1) \quad \begin{aligned} \min & \text{Tr}(\mathbf{X}^T \mathbf{A}_0 \mathbf{X}) + 2\text{Tr}(\mathbf{B}_0^T \mathbf{X}) + c_0 \\ \text{s.t.} & \text{Tr}(\mathbf{X}^T \mathbf{A}_i \mathbf{X}) + 2\text{Tr}(\mathbf{B}_i^T \mathbf{X}) + c_i \leq \alpha_i, i \in \mathcal{I}, \\ & \text{Tr}(\mathbf{X}^T \mathbf{A}_j \mathbf{X}) + 2\text{Tr}(\mathbf{B}_j^T \mathbf{X}) + c_j = \alpha_j, j \in \mathcal{E}, \\ & \mathbf{X} \in \mathbb{R}^{n \times r}, \end{aligned}$$

with $\mathbf{A}_i = \mathbf{A}_i^T \in \mathbb{R}^{n \times n}$, $\mathbf{B}_i \in \mathbb{R}^{n \times r}$, $\alpha_i, c_i \in \mathbb{R}$, $i \in \{0\} \cup \mathcal{I} \cup \mathcal{E}$. Problems of the above type arise naturally in several applications such as robust least squares [9], and in problems involving orthogonal constraints such as the orthogonal procrustes problem [17] (see the discussion in section 2).

Problem (1) is called a *quadratic matrix programming* (QMP) problem of order r . Correspondingly, the objective and constraint functions are called *quadratic matrix (QM) functions*. It can be shown that every QM function is in particular a quadratic function with nr variables; see the discussion in section 2.1. Thus, the family of QMP problems is a special case of quadratically constrained quadratic programming (QCQP) problems. However, it is worthwhile to study these problems independently since, as we shall see, they enjoy stronger results than those currently known for the general QCQP problem. For example, we will establish strong duality results for QMP problems with at most r constraints (see section 3.2).

Strong duality is known to hold for only a few classes of nonconvex QCQP. The simplest and best-known example is the trust region problem, which consists of minimizing an indefinite quadratic function over a ball and admits an exact semidefinite relaxation (SDR); see [13, 8]. Extensions of this problem were considered in

*Received by the editors December 21, 2005; accepted for publication (in revised form) June 30, 2006; published electronically January 12, 2007.

<http://www.siam.org/journals/siopt/17-4/64816.html>

†Department of Industrial Engineering and Management, Technion—Israel Institute of Technology, Haifa 32000, Israel (becka@ie.technion.ac.il).

[12, 18, 5, 16]. In general these results cannot be extended to QCQP problems involving two constraints [19, 20]. An exception is the case in which all the functions involved (objective plus two constraints) are homogenous quadratic functions. In this case, it was proven in [19] that under mild conditions the semidefinite relaxation is tight. Another interesting tractable class of QCQP problems was considered in [1] in the context of quadratic problems with orthogonal constraints.

In this paper strong duality/tightness of the SDR is shown to hold for the class of QMP problems of order r with at most r constraints. In section 3 we construct an SDR and dual formulations for the QMP problem originating from a homogenization procedure specially devised to QMP problems. Using the SDR formulation combined with known results on the existence of low-rank solutions of semidefinite programs [3, 2, 14, 15], the strong duality result is shown to follow. Moreover, an algorithm for extracting a solution to the QMP problem from its associated SDR is described. In section 4 an alternative SDR and dual construction are discussed. These constructions stem from the standard construction of SDR and dual for QCQP problems. Using a result on the equivalence of two linear matrix inequality (LMI) representations of the claim on nonnegativity of a QM function, we are able to prove that the two SDR and dual formulations are equivalent. Finally, in section 5 we present an application of our results in the field of robust optimization.

Notation. For simplicity, instead of inf/sup we use min/max; however, this does not mean that we assume that the optimum is attained and/or finite. Vectors are denoted by boldface lowercase letters, e.g., \mathbf{y} , and matrices by boldface uppercase letters, e.g., \mathbf{A} . For two matrices \mathbf{A} and \mathbf{B} , $\mathbf{A} \succ \mathbf{B}$ ($\mathbf{A} \succeq \mathbf{B}$) means that $\mathbf{A} - \mathbf{B}$ is positive definite (semidefinite). $\mathcal{S}^n = \{\mathbf{A} \in \mathbb{R}^{n \times n} : \mathbf{A} = \mathbf{A}^T\}$ is the set of symmetric $n \times n$ matrices, and $\mathcal{S}_+^n = \{\mathbf{A} \in \mathbb{R}^{n \times n} : \mathbf{A} \succeq \mathbf{0}\}$ is the set all real $n \times n$ symmetric positive semidefinite matrices. $\mathbf{0}_{n \times m}$ is the $n \times m$ matrix of zeros, and \mathbf{I}_r is the $r \times r$ identity matrix. For a matrix \mathbf{M} , $\text{vec}(\mathbf{M})$ denotes the vector obtained by stacking the columns of \mathbf{M} . For a square matrix \mathbf{U} , $[\mathbf{U}]_r$ denotes the southeast $r \times r$ submatrix of \mathbf{U} ; i.e., if $\mathbf{U} = (u_{ij})_{i,j=1}^{n+r}$, then $[\mathbf{U}]_r = (u_{ij})_{i,j=n+1}^{n+r}$. For two matrices \mathbf{A} and \mathbf{B} , $\mathbf{A} \otimes \mathbf{B}$ denotes the corresponding Kronecker product. \mathbf{E}_{ij}^r is the $r \times r$ matrix with 1 at the ij th component and 0 elsewhere, and δ_{ij} is the Kronecker delta, i.e., $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$. The value of the optimal objective function of an optimization problem

$$(P) : \min\{f(\mathbf{x}) : \mathbf{x} \in C\}$$

is denoted by $\text{val}(P)$. The optimization problem (P) is called *bounded below* if the minimum is finite, and termed *solvable* in the case where the minimum is finite and attained (similar definitions for maximum problems). We follow the MATLAB convention and use “;” for adjoining scalars, vectors, or matrices in a column. We also use some standard abbreviations such as SDP (semidefinite programming), LMI (linear matrix inequality), SDR (semidefinite relaxation), and QCQP (quadratically constrained quadratic programming), and some nonstandard abbreviations such as QM (quadratic matrix) and QMP (quadratic matrix programming).

2. Quadratic matrix problems.

2.1. Quadratic matrix functions: Definition and basic properties. We begin by recalling that a quadratic function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is a function of the form

$$(2) \quad g(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + 2\mathbf{b}^T \mathbf{x} + c,$$

where $\mathbf{A} \in \mathcal{S}^n$, $\mathbf{b} \in \mathbb{R}^n$, and $c \in \mathbb{R}$. We will also use the term “quadratic vector function” instead of “quadratic function” to distinguish it from the term “quadratic matrix function” defined below.

A *quadratic matrix (QM) function* of order r is a function $f : \mathbb{R}^{n \times r} \rightarrow \mathbb{R}$ of the form

$$(3) \quad f(\mathbf{X}) = \text{Tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) + 2\text{Tr}(\mathbf{B}^T \mathbf{X}) + c, \quad \mathbf{X} \in \mathbb{R}^{n \times r},$$

where $\mathbf{A} \in \mathcal{S}^n$, $\mathbf{B} \in \mathbb{R}^{n \times r}$, and $c \in \mathbb{R}$. If $\mathbf{B} = \mathbf{0}_{n \times r}$, $c = 0$, then f is called a *homogenous QM function* or a *QM form*. We note that every quadratic vector function is a QM function of order one. The opposite statement is also true: every QM function is in particular a quadratic vector function. Indeed, the function f from (3) can be written as follows:

$$(4) \quad f(\mathbf{X}) = f^V(\text{vec}(\mathbf{X})),$$

where $f^V : \mathbb{R}^{nr} \rightarrow \mathbb{R}$ is defined by

$$(5) \quad f^V(\mathbf{z}) = \mathbf{z}^T (\mathbf{I}_r \otimes \mathbf{A}) \mathbf{z} + 2\text{vec}(\mathbf{B})^T \mathbf{z} + c.$$

The function f^V is called the *vectorized function of f* . From the above relation we can immediately deduce that f is (strictly) convex if and only if $\mathbf{A} \succeq \mathbf{0}$ ($\mathbf{A} \succ \mathbf{0}$).¹

2.2. QM problems. Our main objective is to study *quadratic matrix programming* (QMP) problems in which the goal is to minimize a QM objective function subject to equality and inequality QM constraints:

$$(6) \quad \begin{aligned} (\text{QMP}) \quad & \min f_0(\mathbf{X}) \\ & \text{s.t. } f_i(\mathbf{X}) \leq \alpha_i, i \in \mathcal{I}, \\ & \quad f_j(\mathbf{X}) = \alpha_j, j \in \mathcal{E}, \\ & \quad \mathbf{X} \in \mathbb{R}^{n \times r}, \end{aligned}$$

where $f_i : \mathbb{R}^{n \times r} \rightarrow \mathbb{R}$, $i \in \mathcal{I} \cup \mathcal{E} \cup \{0\}$, are QM functions of order r given by

$$f_i(\mathbf{X}) = \text{Tr}(\mathbf{X}^T \mathbf{A}_i \mathbf{X}) + 2\text{Tr}(\mathbf{B}_i^T \mathbf{X}) + c_i, \quad \mathbf{X} \in \mathbb{R}^{n \times r},$$

with $\mathbf{A}_i \in \mathcal{S}^n$, $\mathbf{B}_i \in \mathbb{R}^{n \times r}$, and $c_i \in \mathbb{R}$, $i \in \{0\} \cup \mathcal{I} \cup \mathcal{E}$. The index sets $\{0\}, \mathcal{I}, \mathcal{E}$ are pairwise disjoint sets of nonnegative integers.

In the case where all the functions f_i , $i \in \mathcal{I} \cup \mathcal{E} \cup \{0\}$, are homogeneous QM functions of order r , the QMP problem (6) is called a *homogenous QMP problem* (of order r). By using the correspondence (4), we can represent the QMP problem as the QCQP problem:

$$(7) \quad \begin{aligned} & \min f_0^V(\mathbf{z}) \\ & \text{s.t. } f_i^V(\mathbf{z}) \leq \alpha_i, i \in \mathcal{I}, \\ & \quad f_j^V(\mathbf{z}) = \alpha_j, j \in \mathcal{E}, \\ & \quad \mathbf{z} \in \mathbb{R}^{nr}, \end{aligned}$$

which will be called the *vectorized QMP problem*.

¹Indeed, $\mathbf{I}_r \otimes \mathbf{A}$ and \mathbf{A} have the same eigenvalues (but with different multiplicities) [10].

The QMP problem appears in several fields of applications. Here we present two examples in which the QMP problem naturally arises.

Example 1. In the *orthogonal procrustes* problem [17] we seek to find a square matrix \mathbf{X} which solves the following optimization problem:

$$\begin{aligned} \min & \|\mathbf{A}\mathbf{X} - \mathbf{B}\|_F^2 \\ \text{s.t.} & \mathbf{X}^T\mathbf{X} = \mathbf{I}_r, \\ & \mathbf{X} \in \mathbb{R}^{r \times r}, \end{aligned}$$

where $\mathbf{A} \in \mathbb{R}^{n \times r}$, $\mathbf{B} \in \mathbb{R}^{n \times r}$. The orthogonal procrustes problem can be rewritten as a QMP problem with r^2 equality constraints:

$$\begin{aligned} \min & \text{Tr}(\mathbf{X}^T \mathbf{A}^T \mathbf{A} \mathbf{X}) - 2\text{Tr}(\mathbf{B}^T \mathbf{A} \mathbf{X}) + \|\mathbf{B}\|_F^2 \\ \text{s.t.} & \text{Tr}(\mathbf{X}^T (\mathbf{E}_{ij}^r + \mathbf{E}_{ij}^r) \mathbf{X}) = 2\delta_{ij}, \quad 1 \leq i, j \leq r, \\ & \mathbf{X} \in \mathbb{R}^{r \times r}. \end{aligned}$$

We note that although the orthogonal procrustes problem can be solved efficiently [17], it is not clear whether the *unbalanced* orthogonal procrustes problem—in which \mathbf{X} is not square—is tractable [7].

Example 2. The *robust least squares* (RLS) problem was introduced and studied in [9, 6].² Consider a linear system $\mathbf{A}\mathbf{x} \approx \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{r \times n}$, $\mathbf{b} \in \mathbb{R}^r$, and $\mathbf{x} \in \mathbb{R}^n$. Assume that the matrix \mathbf{A} is not fixed but rather given by a family of matrices³ $\mathbf{A} + \mathbf{\Delta}^T$, where \mathbf{A} is a known nominal value and $\mathbf{\Delta} \in \mathbb{R}^{n \times r}$ is an unknown perturbation matrix known to reside in a compact uncertainty set \mathcal{U} . The RLS approach to this problem is to seek a vector $\mathbf{x} \in \mathbb{R}^n$ that minimizes the worst case data error with respect to all possible values of $\mathbf{\Delta} \in \mathcal{U}$:

$$(8) \quad \min_{\mathbf{x}} \max_{\mathbf{\Delta} \in \mathcal{U}} \|\mathbf{b} - (\mathbf{A} + \mathbf{\Delta}^T)\mathbf{x}\|^2.$$

Now, by making some simple algebraic manipulation, we can rewrite the objective function in (8) as

$$\|\mathbf{b} - (\mathbf{A} + \mathbf{\Delta}^T)\mathbf{x}\|^2 = \text{Tr}(\mathbf{\Delta}^T \mathbf{x} \mathbf{x}^T \mathbf{\Delta}) + 2\text{Tr}((\mathbf{b} - \mathbf{A}\mathbf{x})\mathbf{x}^T \mathbf{\Delta}) + \text{Tr}((\mathbf{b} - \mathbf{A}\mathbf{x})(\mathbf{b} - \mathbf{A}\mathbf{x})^T),$$

so that the inner maximization problem in (8) takes the following form:

$$(9) \quad \max\{\text{Tr}(\mathbf{\Delta}^T \mathbf{Q} \mathbf{\Delta}) + 2\text{Tr}(\mathbf{F}^T \mathbf{\Delta}) + c : \mathbf{\Delta} \in \mathcal{U}\},$$

where \mathbf{Q} , \mathbf{F} , and c depend on \mathbf{x} and are given by

$$(10) \quad \mathbf{Q} = \mathbf{x}\mathbf{x}^T \in \mathcal{S}^n, \quad \mathbf{F} = \mathbf{x}(\mathbf{b} - \mathbf{A}\mathbf{x})^T \in \mathbb{R}^{n \times r}, \quad c = \|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 \in \mathbb{R}.$$

In [9] the uncertainty set \mathcal{U} was chosen to be a simple Frobenius norm constraint, i.e.,

$$\mathcal{U} = \{\mathbf{\Delta} \in \mathbb{R}^{n \times r} : \text{Tr}(\mathbf{\Delta}^T \mathbf{\Delta}) \leq \rho\}.$$

²Here we study the unstructured case.

³The perturbation matrix appears in a transpose form so that the derived QM function will have the form (3). Furthermore, for the sake of simplicity we do not consider uncertainties in the RHS vector \mathbf{b} , although such uncertainties can be incorporated into our analysis in a straightforward manner.

The inner maximization problem (9) with the above choice of \mathcal{U} is a QMP problem of order r with a single inequality constraint.

The fact that the uncertainty set \mathcal{U} was given in [9] by a *single* quadratic constraint was a crucial element in establishing the tractability of the RLS problem. In fact, it is well known that in the structured case, the inner maximization problem of the RLS problem becomes NP-hard when the uncertainty set is given by an intersection of ellipsoids. Nonetheless, in section 5, using the results developed in sections 3 and 4, we will show that more complicated choices of \mathcal{U} can be considered. In particular, we will prove in section 5 that the RLS problem remains tractable in the case where \mathcal{U} is given by a set of at most r QM inequality constraints. The latter form of the uncertainty set can model, for example, the situation where each column of the perturbation matrix Δ^T has a *separate* norm constraint.

3. Semidefinite relaxations of the QMP problem and strong duality results. We begin by constructing an SDR for the QMP problem. A natural approach for constructing such an SDR is to consider the SDR of the vectorized problem (7) (recall that problem (7) is a (QCQP)). However, this approach, which is discussed in detail in section 4, does not seem to offer useful theoretical insights into questions such as strong duality/tightness of SDR. For that reason we construct a *new* scheme, specifically devised to obtain an SDR for QMP problems (see section 3.1). Using the derived SDR, we will show in section 3.2 that, under some mild conditions, strong duality holds for QMP problems of order r with at most r constraints.

3.1. An SDR of the QMP problem. Recall that the homogenized version a quadratic vector function g given by (2) is the quadratic form $g^H : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ defined by

$$(11) \quad g^H(\mathbf{x}; t) = \mathbf{x}^T \mathbf{A} \mathbf{x} + 2\mathbf{b}^T \mathbf{x} t + ct^2.$$

The matrix associated with the quadratic form g^H is denoted by

$$(12) \quad M(g) = \begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & c \end{pmatrix}.$$

We consider the following generalization of the above homogenization procedure to QM functions of order r : let f be the QM function given by (3); the *homogenized QM function* is denoted by $f^H : \mathbb{R}^{(n+r) \times r} \rightarrow \mathbb{R}$ and given by

$$(13) \quad f^H(\mathbf{Y}; \mathbf{Z}) \equiv \text{Tr}(\mathbf{Y}^T \mathbf{A} \mathbf{Y}) + 2\text{Tr}(\mathbf{Z}^T \mathbf{B}^T \mathbf{Y}) + \frac{c}{r} \text{Tr}(\mathbf{Z}^T \mathbf{Z}), \quad \mathbf{Y} \in \mathbb{R}^{n \times r}, \mathbf{Z} \in \mathbb{R}^{r \times r},$$

which is a homogenous QM function of order r corresponding to the matrix

$$(14) \quad M(f) \equiv \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \frac{c}{r} \mathbf{I}_r \end{pmatrix}.$$

In the case $r = 1$, definitions (13) and (14) coincide with the definitions of the homogenization of a quadratic function (11) and its associated matrix (12), respectively. The operator M will be used throughout the paper.

The homogenous function f^H satisfies the following easily verifiable properties, which will become useful in what follows:

$$(15) \quad f^H(\mathbf{Y}; \mathbf{I}_r) = f(\mathbf{Y}) \quad \text{for every } \mathbf{Y} \in \mathbb{R}^{n \times r},$$

$$(16) \quad f^H(\mathbf{Y}; \mathbf{Z}) = f(\mathbf{Y} \mathbf{Z}^T) \quad \text{for every } \mathbf{Y} \in \mathbb{R}^{n \times r}, \mathbf{Z} \in \mathbb{R}^{r \times r} \text{ such that } \mathbf{Z}^T \mathbf{Z} = \mathbf{I}_r.$$

Using the above homogenization procedure for QM functions, we are able to construct (see Lemma 3.1 below) a homogeneous QMP problem of order r , equivalent to the (nonhomogeneous) QMP problem (6).

LEMMA 3.1. *Consider the following homogenized version of the QMP problem (6):*

$$(17) \quad \begin{aligned} \min f_0^H(\mathbf{Y}; \mathbf{Z}) \\ \text{s.t. } f_i^H(\mathbf{Y}; \mathbf{Z}) \leq \alpha_i, i \in I, \\ f_j^H(\mathbf{Y}; \mathbf{Z}) = \alpha_j, j \in \mathcal{E}, \\ \psi_{ij}(\mathbf{Y}; \mathbf{Z}) = 2\delta_{ij}, 1 \leq i \leq j \leq r, \\ \mathbf{Y} \in \mathbb{R}^{n \times r}, \mathbf{Z} \in \mathbb{R}^{r \times r}, \end{aligned}$$

where $\psi_{ij}(\mathbf{Y}; \mathbf{Z}) = \text{Tr}(\mathbf{Z}^T(\mathbf{E}_{ij}^r + \mathbf{E}_{ji}^r)\mathbf{Z})$ and δ_{ij} is the Kronecker delta.

1. Suppose that the QMP problem (6) is solvable, and let \mathbf{X}^* be an optimal solution of (QMP). Then problem (17) is solvable, $(\mathbf{X}^*; \mathbf{I}_r)$ is an optimal solution of (17), and $\text{val}(\text{QMP}) = \text{val}(17)$.
2. Suppose that problem (17) is solvable, and let $(\mathbf{Y}^*; \mathbf{Z}^*)$ be an optimal solution of (17). Then problem (QMP) is solvable, $\mathbf{X}^* = \mathbf{Y}^*(\mathbf{Z}^*)^T$ is an optimal solution of (QMP), and $\text{val}(\text{QMP}) = \text{val}(17)$.

Proof. First note that the system of equalities

$$\text{Tr}(\mathbf{Z}^T(\mathbf{E}_{ij}^r + \mathbf{E}_{ji}^r)\mathbf{Z}) = 2\delta_{ij}, \quad 1 \leq i \leq j \leq r,$$

can be written as

$$\text{Tr}((\mathbf{E}_{ij}^r + \mathbf{E}_{ji}^r)\mathbf{Z}\mathbf{Z}^T) = 2\delta_{ij}, \quad 1 \leq i \leq j \leq r,$$

which, by using the symmetry of the matrix $\mathbf{Z}\mathbf{Z}^T$, is equivalent to

$$\mathbf{Z}^T\mathbf{Z} = \mathbf{Z}\mathbf{Z}^T = \mathbf{I}_r.$$

1. Let \mathbf{X}^* be an optimal solution of (QMP). For every $(\mathbf{Y}; \mathbf{Z})$, $(\mathbf{Y} \in \mathbb{R}^{n \times r}, \mathbf{Z} \in \mathbb{R}^{r \times r})$ in the feasible set of (17) (and in particular $\mathbf{Z}^T\mathbf{Z} = \mathbf{I}_r$) we have

$$f_0^H(\mathbf{Y}, \mathbf{Z}) \stackrel{(16)}{=} f_0(\mathbf{Y}\mathbf{Z}^T) \geq f_0(\mathbf{X}^*) \stackrel{(15)}{=} f_0^H(\mathbf{X}^*; \mathbf{I}_r).$$

Therefore, $(\mathbf{X}^*; \mathbf{I}_r)$ is an optimal solution of (17) and $\text{val}(\text{QMP}) = \text{val}(17)$.

2. Let $(\mathbf{Y}^*; \mathbf{Z}^*)$, $(\mathbf{Y} \in \mathbb{R}^{n \times r}, \mathbf{Z} \in \mathbb{R}^{r \times r})$ be an optimal solution of (17), and set $\mathbf{X}^* = \mathbf{Y}^*(\mathbf{Z}^*)^T$. Then for every $\mathbf{X} \in \mathbb{R}^{n \times r}$ which is in the feasible set of (QMP) we have

$$f_0(\mathbf{X}) \stackrel{(15)}{=} f_0^H(\mathbf{X}; \mathbf{I}) \geq f_0^H(\mathbf{Y}^*; \mathbf{Z}^*) \stackrel{(16)}{=} f_0(\mathbf{Y}^*(\mathbf{Z}^*)^T) = f_0(\mathbf{X}^*),$$

and thus \mathbf{X}^* is an optimal solution of (QMP) and $\text{val}(\text{QMP}) = \text{val}(17)$. \square

COROLLARY 3.2. *The QMP problem (6) is solvable if and only if problem (17) is solvable, and in that case $\text{val}(\text{QMP}) = \text{val}(17)$.*

We will now exploit the homogenized QMP problem (17) in order to formulate a semidefinite relaxation. By denoting $W = (\mathbf{Y}; \mathbf{Z}) \in \mathbb{R}^{(n+r) \times r}$, we conclude that

problem (17) can be written as

$$\begin{aligned} & \min \operatorname{Tr}(\mathbf{M}(f_0)\mathbf{W}\mathbf{W}^T) \\ & \text{s.t. } \operatorname{Tr}(\mathbf{M}(f_i)\mathbf{W}\mathbf{W}^T) \leq \alpha_i, i \in I, \\ & \quad \operatorname{Tr}(\mathbf{M}(f_j)\mathbf{W}\mathbf{W}^T) = \alpha_j, j \in \mathcal{E}, \\ & \quad \operatorname{Tr}(\mathbf{N}_{ij}\mathbf{W}\mathbf{W}^T) = 2\delta_{ij}, 1 \leq i \leq j \leq r, \\ & \quad \mathbf{W} \in \mathbb{R}^{(n+r) \times r}, \end{aligned}$$

where the operator \mathbf{M} is defined in (14) and

$$\mathbf{N}_{ij} = \begin{pmatrix} \mathbf{0}_{n \times n} & \mathbf{0}_{n \times r} \\ \mathbf{0}_{r \times n} & \mathbf{E}_{ij}^r + \mathbf{E}_{ji}^r \end{pmatrix}, \quad 1 \leq i \leq j \leq r.$$

Making the change of variables $\mathbf{U} = \mathbf{W}\mathbf{W}^T \in \mathcal{S}_+^{n+r}$, we conclude that problem (17) can be equivalently written as

$$\begin{aligned} & \min \operatorname{Tr}(\mathbf{M}(f_0)\mathbf{U}) \\ & \text{s.t. } \operatorname{Tr}(\mathbf{M}(f_i)\mathbf{U}) \leq \alpha_i, i \in I, \\ & \quad \operatorname{Tr}(\mathbf{M}(f_j)\mathbf{U}) = \alpha_j, j \in \mathcal{E}, \\ & \quad \operatorname{Tr}(\mathbf{N}_{ij}\mathbf{U}) = 2\delta_{ij}, 1 \leq i \leq j \leq r, \\ & \quad \mathbf{U} \in \mathcal{S}_+^{n+r}, \operatorname{rank}(\mathbf{U}) \leq r. \end{aligned}$$

Omitting the “hard” constraint $\operatorname{rank}(\mathbf{U}) \leq r$, we finally arrive at the following SDR of the QMP problem (6):

$$\begin{aligned} (18) \quad & \text{(SDRM)} \quad \min \operatorname{Tr}(\mathbf{M}(f_0)\mathbf{U}) \\ & \text{s.t. } \operatorname{Tr}(\mathbf{M}(f_i)\mathbf{U}) \leq \alpha_i, i \in I, \\ & \quad \operatorname{Tr}(\mathbf{M}(f_j)\mathbf{U}) = \alpha_j, j \in \mathcal{E}, \\ & \quad \operatorname{Tr}(\mathbf{N}_{ij}\mathbf{U}) = 2\delta_{ij}, 1 \leq i \leq j \leq r, \\ & \quad \mathbf{U} \in \mathcal{S}_+^{n+r}. \end{aligned}$$

The dual problem to the SDR problem (SDRM) is given by

$$\begin{aligned} (19) \quad & \text{(DM)} \quad \max_{\lambda_i, \Phi} - \sum_{i \in \mathcal{I} \cup \mathcal{E}} \lambda_i \alpha_i - \operatorname{Tr}(\Phi) \\ & \text{s.t. } \mathbf{M}(f_0) + \sum_{i \in \mathcal{I} \cup \mathcal{E}} \lambda_i \mathbf{M}(f_i) + \begin{pmatrix} \mathbf{0}_{n \times n} & \mathbf{0}_{n \times r} \\ \mathbf{0}_{r \times n} & \Phi \end{pmatrix} \succeq \mathbf{0}, \\ & \quad \Phi \in \mathcal{S}^r, \\ & \quad \lambda_i \geq 0, i \in \mathcal{I}. \end{aligned}$$

The symmetric matrix $\Phi = (\phi_{ij})_{i,j=1}^r$ contains the Lagrange multipliers associated with the equality constraints $\operatorname{Tr}(\mathbf{N}_{ij}\mathbf{U}) = 2\delta_{ij}$. Specifically, for every $1 \leq i \leq r$, $\frac{1}{2}\phi_{ii}$ is the multiplier corresponding to the constraint $\operatorname{Tr}(\mathbf{N}_{ii}\mathbf{U}) = 2$, and ϕ_{ij} ($= \phi_{ji}$) is the multiplier associated with $\operatorname{Tr}(\mathbf{N}_{ij}\mathbf{U}) = 0$ for $1 \leq i < j \leq r$. By the conic duality theorem [4] it follows that if (DM) is strictly feasible and bounded above, then (SDRM) is solvable and $\operatorname{val}(\text{SDRM}) = \operatorname{val}(\text{DM})$. For that reason we seek to find a simple condition under which (DM) is strictly feasible. The following lemma establishes such a condition.

LEMMA 3.3. *Suppose that the following condition is satisfied:*

$$(20) \quad \exists \gamma_i \in \mathbb{R}, i \in \mathcal{I} \cup \mathcal{E}, \text{ for which } \gamma_i \geq 0, i \in \mathcal{I}, \text{ such that } \mathbf{A}_0 + \sum_{i \in \mathcal{I} \cup \mathcal{E}} \gamma_i \mathbf{A}_i \succ \mathbf{0}.$$

Then problem (DM) is strictly feasible.

Proof. Let $\gamma_i \in \mathbb{R}, i \in \mathcal{I} \cup \mathcal{E}$, be numbers satisfying (20), and let $\epsilon > 0$ be a small enough number for which $\mathbf{A}_0 + \sum_{i \in \mathcal{I} \cup \mathcal{E}} (\gamma_i + \epsilon) \mathbf{A}_i \succ \mathbf{0}$. Define $\tilde{\gamma}_i \equiv \gamma_i + \epsilon$. Evidently, $\tilde{\gamma}_i > 0$ for $i \in \mathcal{I}$. Now, for every symmetric $r \times r$ matrix Φ we have

$$(21) \quad \begin{aligned} & \mathbf{M}(f_0) + \sum_{i \in \mathcal{I} \cup \mathcal{E}} \tilde{\gamma}_i \mathbf{M}(f_i) + \begin{pmatrix} \mathbf{0}_{n \times n} & \mathbf{0}_{n \times r} \\ \mathbf{0}_{r \times n} & \Phi \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A}_0 + \sum \tilde{\gamma}_i \mathbf{A}_i & \mathbf{B}_0 + \sum \tilde{\gamma}_i \mathbf{B}_i \\ (\mathbf{B}_0 + \sum \tilde{\gamma}_i \mathbf{B}_i)^T & \frac{1}{r} (c_0 + \sum \tilde{\gamma}_i c_i) \mathbf{I}_r + \Phi \end{pmatrix}, \end{aligned}$$

where all the summations are over $i \in \mathcal{I} \cup \mathcal{E}$. Since $\mathbf{A}_0 + \sum \tilde{\gamma}_i \mathbf{A}_i \succ \mathbf{0}$, then by the Schur complement, the matrix on the RHS of (21) is positive definite if and only if

$$\Phi \succ \left(\mathbf{B}_0 + \sum \tilde{\gamma}_i \mathbf{B}_i \right)^T \left(\mathbf{A}_0 + \sum \tilde{\gamma}_i \mathbf{A}_i \right)^{-1} \left(\mathbf{B}_0 + \sum \tilde{\gamma}_i \mathbf{B}_i \right) - \frac{1}{r} \left(c_0 + \sum \tilde{\gamma}_i c_i \right) \mathbf{I}_r.$$

Let $\tilde{\Phi} \in \mathcal{S}^r$ be an arbitrary matrix satisfying the latter LMI. Thus, for $\lambda_i = \tilde{\gamma}_i, i \in \mathcal{I} \cup \mathcal{E}$, and $\Phi = \tilde{\Phi}$ we have that all the inequalities in (19) (regular and generalized) are strictly satisfied. \square

Remark 3.1. Conditions similar to (20) are very common in the analysis of QCQP problems; see, e.g., [5, 18, 12, 19, 16]. This condition is automatically satisfied when at least one of the constraints or the objective function is strictly convex (see also [19, Proposition 2.1]).

3.2. Tightness of the SDR of the QMP problem. In this section we will show that, under some mild conditions, QMP problems of order r with at most r constraints have a tight SDR, and that strong duality holds. To show this, we need to verify that problem (SDRM) possesses a solution with rank smaller than or equal to r . This prompts us to consider questions concerning the existence of low-rank solutions to SDP problems—a subject extensively studied by Pataki [14, 15] and Barvinok [2, 3]; see also [11] for related results concerning the convexity of the image of several homogenous QMs.

Let us consider a general-form SDP problem:

$$(22) \quad \begin{aligned} & \min \text{Tr}(\mathbf{C}_0 \mathbf{U}) \\ & \text{s.t. } \text{Tr}(\mathbf{C}_i \mathbf{U}) \leq \alpha_i, i \in \mathcal{I}_1, \\ & \quad \text{Tr}(\mathbf{C}_j \mathbf{U}) = \alpha_j, j \in \mathcal{E}_1, \\ & \quad \mathbf{U} \in \mathcal{S}_+^n, \end{aligned}$$

where \mathcal{I}_1 and \mathcal{E}_1 are disjoint index sets, $\mathbf{C}_i \in \mathcal{S}^n, i \in \{0\} \cup \mathcal{I}_1 \cup \mathcal{E}_1$, and $\alpha_i \in \mathbb{R}, i \in \mathcal{I}_1 \cup \mathcal{E}_1$. Pataki showed [15] that if the number of constraints is smaller than an upper bound which is a certain quadratic function of r , then there exists a solution with rank no larger than r (see Theorem 3.4 below). The proof of this result is constructive and is based on a simple rank reduction procedure⁴ for finding extreme points of convex

⁴The SDP considered in [15] consists only of inequality constraints. However, the same analysis establishes the validity of Theorem 3.4.

sets of the form $\mathcal{S}_+^n \cap \mathcal{A}$, where \mathcal{A} is an affine space. For the sake of completeness, and since the rank reduction procedure is a subroutine of the algorithm for solving the QMP problem, we recall both the claim (Theorem 3.4 below) and the rank reduction procedure (see Algorithm RED in the appendix).

THEOREM 3.4 (see [15]). *Suppose that problem (22) is solvable and that $|\mathcal{I}_1| + |\mathcal{E}_1| \leq \binom{r+2}{2}$, where r is a positive integer. Then problem (22) has a solution \mathbf{X}^* for which $\text{rank}(\mathbf{X}^*) \leq r$.*

Proof. Let \mathbf{X}_0^* be an optimal solution of problem (22). Apply Algorithm RED (see the appendix) with input \mathbf{X}_0^* and obtain an optimal solution \mathbf{X}^* with $\text{rank}(\mathbf{X}^*) \leq r$. \square

Equipped with the latter result, we are now able to show that QMP problems of order r with at most r constraints possess a tight SDR under some mild conditions.

THEOREM 3.5 (tight SDR for the QMP problem). *If problem (SDRM) is solvable and $|\mathcal{I}| + |\mathcal{E}| \leq r$, then problem (QMP) is solvable and $\text{val}(\text{SDRM}) = \text{val}(\text{QMP})$.*

Proof. It is sufficient to show that problem (SDRM) has a solution with rank smaller than or equal to r . The number of constraints in (SDRM) is equal to $|\mathcal{I}| + |\mathcal{E}| + \binom{r+1}{2}$, where the last term stands for the number of pairs (i, j) for which $1 \leq i \leq j \leq r$. Thus, using $|\mathcal{I}| + |\mathcal{E}| \leq r$, we conclude that the number of constraints in (SDRM) is bounded above by

$$r + \binom{r+1}{2} = \binom{r+2}{2} - 1.$$

Invoking Theorem 3.4, the result follows. \square

As a conclusion from the conic duality theorem [4] we can now deduce the following corollary that guarantees tightness of the SDR and strong duality under the conditions that the QMP problem (6) is feasible and that condition (20) is valid.

COROLLARY 3.6 (strong duality for QMP problems). *Consider the QMP problem (6) with $|\mathcal{I}| + |\mathcal{E}| \leq r$, its semidefinite relaxation (SDRM) (problem (18)) and its dual (DM) (problem (19)). Suppose that condition (20) holds true and that the QMP problem is feasible. Then problems (QMP) and (SDRM) are solvable and $\text{val}(\text{QMP}) = \text{val}(\text{SDRM}) = \text{val}(\text{DM})$.*

Proof. By Lemma 3.3, the validity of condition (20) implies that the dual problem (DM) is strictly feasible. Moreover, since the primal SDP problem (SDRM) is feasible, it follows that the dual problem (DM) is bounded above. Thus, by the conic duality theorem [4], we conclude that problem (SDRM) is solvable and that $\text{val}(\text{SDRM}) = \text{val}(\text{DM})$. Since problem (SDRM) is solvable we conclude, by Theorem 3.5, that $\text{val}(\text{QMP}) = \text{val}(\text{SDRM})$. \square

Remark 3.2. In the special case $r = 1$, Corollary 3.6 recovers the well-known strong duality/tightness of SDR results for QCQPs with a single quadratic constraint (see, e.g., [12, 5, 18, 16]).

It is interesting to note that we can also describe an algorithm for extracting the solution of a QMP problem (satisfying the condition in Corollary 3.6) from its SDR, which is based on the rank reduction algorithm of [15], as follows.

ALGORITHM SOL-QMP.

Step 1. Solve the SDP problem (SDRM) and obtain an optimal solution $\mathbf{U}^* \in \mathcal{S}_+^{n+r}$.

Step 2. Invoke Algorithm RED (see the appendix) with input \mathbf{U}^* , and produce an optimal solution $\mathbf{U}_1^* \in \mathcal{S}_+^{n+r}$ for which $\text{rank}(\mathbf{U}_1^*) \leq r$.

Step 3. Calculate a decomposition: $\mathbf{U}_1^* = \mathbf{W}\mathbf{W}^T$, where $\mathbf{W} \in \mathbb{R}^{(n+r) \times r}$.

Step 4. Let $\mathbf{W} = (\mathbf{Y}; \mathbf{Z})$, where $\mathbf{Y} \in \mathbb{R}^{n \times r}$ and $\mathbf{Z} \in \mathbb{R}^{r \times r}$. Return an optimal solution $\mathbf{X}^* = \mathbf{Y}\mathbf{Z}^T$ to the QMP problem.

4. The vectorized semidefinite relaxation and dual of the QMP problem. In the previous section we considered a semidefinite relaxation that was based on a homogenization procedure specifically designed for QM functions. In this section we examine an alternative (and natural) approach in which we begin by transforming the problem into a “standard” QCQP and then use the usual relaxation technique. This approach produces the *vectorized SDR* and *vectorized dual* problems. We will prove that the two constructions are equivalent in some sense. In establishing this result we rely on the tight SDR result of section 3 and a result on two LMI representations of the property of nonnegativity of a QM function over $\mathbb{R}^{n \times r}$.

Our alternative SDR is constructed by following two steps.

Step 1. Transform the QMP problem (6) into the vectorized QMP problem (7).

Step 2. Formulate the corresponding SDR of the homogenized problem (7):

$$\begin{aligned}
 (SDRV) \quad & \min \operatorname{Tr}(\mathbf{M}(f_0^V)\mathbf{Z}) \\
 (23) \quad & \text{s.t. } \operatorname{Tr}(\mathbf{M}(f_i^V)\mathbf{Z}) \leq \alpha_i, i \in \mathcal{I}, \\
 & \operatorname{Tr}(\mathbf{M}(f_j^V)\mathbf{Z}) = \alpha_j, j \in \mathcal{E}, \\
 & Z_{nr+1, nr+1} = 1, \\
 & \mathbf{Z} \in \mathcal{S}_+^{nr+1}
 \end{aligned}$$

(recall that, since f_i^V is a QM function of order one, $\mathbf{M}(f_i^V) \equiv \begin{pmatrix} \mathbf{I}_r \otimes \mathbf{A}_i & \operatorname{vec}(\mathbf{B}_i) \\ \operatorname{vec}(\mathbf{B}_i)^T & c_i \end{pmatrix}$).

Problem (SDRV) is an SDP problem, and its dual is given by

$$\begin{aligned}
 (DV) \quad & \max - \sum_{i \in \mathcal{I} \cup \mathcal{E}} \lambda_i \alpha_i - t \\
 (24) \quad & \text{s.t. } \mathbf{M}(f_0^V) + \sum_{i \in \mathcal{I} \cup \mathcal{E}} \lambda_i \mathbf{M}(f_i^V) + t \begin{pmatrix} \mathbf{0}_{nr, nr} & \mathbf{0}_{nr, 1} \\ \mathbf{0}_{1, nr} & 1 \end{pmatrix} \succeq \mathbf{0}, \\
 & \lambda_i \geq 0, i \in \mathcal{I}.
 \end{aligned}$$

It can be shown that problem (DV) is in fact a Lagrangian dual of the QMP problem (6), and therefore the SDR (SDRV) can be interpreted as a bidual (i.e., dual of the dual) of the primal QMP problem. Problems (SDRV) and (DV) are called the *vectorized semidefinite relaxation* and *dual of the QMP problem* (respectively).

The pair of problems (SDRM)/(SDRV) and (DM)/(DV) seem quite different both with respect to the number of variables and the sizes of the related matrices. However, we will show in what follows (cf. Theorem 4.3) that these pairs of problems are equivalent in some sense.

Lemma 4.2 below presents two different LMI characterizations of the nonnegativity of a QM function over the entire space. This lemma is a key ingredient in proving the equivalence between the different dual/SDR problems. The proof of Lemma 4.2 relies on the following well-known result.

LEMMA 4.1 (see [4, p. 163]). *A quadratic inequality with a (symmetric) $n \times n$ matrix \mathbf{A} ,*

$$\mathbf{x}^T \mathbf{A} \mathbf{x} + 2\mathbf{b}^T \mathbf{x} + c \geq 0,$$

is valid for all $\mathbf{x} \in \mathbb{R}^n$ if and only if

$$\begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & c \end{pmatrix} \succeq \mathbf{0}.$$

LEMMA 4.2. *Let f be a QM function given in (3). Then the following three statements are equivalent:*

- (i) $f(\mathbf{X}) \geq 0$ for every $\mathbf{X} \in \mathbb{R}^{n \times r}$.
- (ii) There exists $\Phi \in \mathcal{S}^r$ for which $\text{Tr}(\Phi) \leq 0$ such that

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \frac{c}{r} \mathbf{I}_r + \Phi \end{pmatrix} \succeq \mathbf{0}.$$

(iii)

$$\begin{pmatrix} \mathbf{I}_r \otimes \mathbf{A} & \text{vec}(\mathbf{B}) \\ \text{vec}(\mathbf{B})^T & c \end{pmatrix} \succeq \mathbf{0}.$$

Proof. (i \Leftrightarrow iii) By (4), the first statement is equivalent to the statement

$$f^V(\mathbf{z}) \geq 0 \text{ for every } \mathbf{z} \in \mathbb{R}^{nr},$$

which, by Lemma 4.1, is the same as the third statement.

(i \Leftrightarrow ii) We begin by showing the following identity between subsets of \mathbb{R} :

$$(25) \quad F = W,$$

where (recall that $[\mathbf{U}]_r$ denotes the southeast $r \times r$ submatrix of \mathbf{U})

$$\begin{aligned} F &= \{f(\mathbf{X}) : \mathbf{X} \in \mathbb{R}^{n \times r}\}, \\ W &= \{\text{Tr}(\mathbf{M}(f)\mathbf{U}) : \mathbf{U} \in \mathcal{S}_+^{n+r}, [\mathbf{U}]_r = \mathbf{I}_r\}. \end{aligned}$$

The inclusion $F \subseteq W$ is clear. We will show that the reverse inclusion ($W \subseteq F$) holds true. Let $\alpha \in W$, and consider the QMP problem

$$(26) \quad \begin{aligned} &\min 0 \\ &\text{s.t. } f(\mathbf{X}) = \alpha, \\ &\quad \mathbf{X} \in \mathbb{R}^{n \times r}. \end{aligned}$$

Note that this is exactly the QMP problem (6) with $r = 1$, $\mathcal{I} = \emptyset$, $\mathcal{E} = \{1\}$, $\alpha_1 = \alpha$, $f_0 \equiv 0$, and $f_1 = f$. The corresponding SDR of the QMP problem (26) is given by

$$(27) \quad \begin{aligned} &\min 0 \\ &\text{s.t. } \text{Tr}(\mathbf{M}(f)\mathbf{U}) = \alpha, \\ &\quad \mathbf{U} \in \mathcal{S}_+^{n+r}, [\mathbf{U}]_r = \mathbf{I}_r. \end{aligned}$$

Since $\alpha \in W$ it follows that problem (27) is solvable (recall that the objective function is identically zero, and hence ‘‘solvability’’ is the same as ‘‘feasibility’’). Invoking Theorem 3.5, we conclude that problem (26) is also feasible. Hence, $\alpha \in F$. The identity $F = W$ implies that statement (i) is the same as

$$(28) \quad \min\{\text{Tr}(\mathbf{M}(f)\mathbf{U}) : \mathbf{U} \in \mathcal{S}_+^{n+r}, [\mathbf{U}]_r = \mathbf{I}_r\} \geq 0.$$

The latter SDP problem is strictly feasible ($\mathbf{U} = \mathbf{I}_{n+r} \succ \mathbf{0}$ is feasible) and bounded below (by zero) and thus, by the conic duality theorem, we conclude that the dual problem, given in this case by

$$\max_{\Phi \in \mathcal{S}^r} \left\{ -\text{Tr}(\Phi) : \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \frac{c}{r} \mathbf{I}_r + \Phi \end{pmatrix} \succeq \mathbf{0} \right\},$$

is solvable and has value equal to the value of the primal problem. Therefore, statement (28) is equivalent to the existence of a symmetric $r \times r$ matrix Φ for which

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \frac{c}{r}\mathbf{I}_r + \Phi \end{pmatrix} \succeq \mathbf{0}$$

and $\text{Tr}(\Phi) \leq 0$. \square

We are now ready to prove the main result of this section, namely, that the values of the two dual problems (DM) and (DV) and the two SDR problems (SDRM) and (SDRV) are all equal to each other under some mild conditions.

THEOREM 4.3. *Consider the SDRs (SDRM) and (SDRV) (problems (18) and (23)) and the dual problems (DM) and (DV) (problems (19) and (24)) of the QMP problem (6). Suppose that condition (20) is satisfied and that (QMP) is feasible. Then (SDRM) and (SDRV) are solvable and*

$$\text{val}(\text{DM}) = \text{val}(\text{DV}) = \text{val}(\text{SDRM}) = \text{val}(\text{SDRV}).$$

Furthermore, if $\{\lambda_i\}_{i \in \mathcal{I} \cup \mathcal{E}}$ and Φ is an optimal solution of (DM), then an optimal solution to (DV) is given by $\{\lambda_i\}_{i \in \mathcal{I} \cup \mathcal{E}}, t$, where $t = \text{Tr}(\Phi)$.

Proof. Since condition (20) is assumed to hold true then, by Lemma 3.3, the dual problem (DM) is strictly feasible, and an argument similar to the one used in the proof of Lemma 3.3 shows that (DV) is also strictly feasible. Thus, by the conic duality Theorem [4], both problems (SDRM) and (SDRV) are solvable, and we have the equality $\text{val}(\text{DM}) = \text{val}(\text{SDRM})$ as well as $\text{val}(\text{DV}) = \text{val}(\text{SDRV})$. We are left with the task of proving that $\text{val}(\text{DM}) = \text{val}(\text{DV})$. Consider the LMI constraint in problem (DV), which can explicitly be written as follows:

$$(29) \quad \begin{pmatrix} \mathbf{I}_r \otimes (\mathbf{A}_0 + \sum \lambda_i \mathbf{A}_i) & \text{vec}(\mathbf{B}_0 + \sum \lambda_i \mathbf{B}_i) \\ \text{vec}(\mathbf{B}_0 + \sum \lambda_i \mathbf{B}_i)^T & c_0 + \sum \lambda_i c_i + t \end{pmatrix} \succeq \mathbf{0},$$

where the summations are over $i \in \mathcal{I} \cup \mathcal{E}$. By the equivalence of the second and third part of Lemma 4.2 we have that the above LMI holds true if and only if there exists $\mathbf{Z} \in \mathcal{S}^r$ such that

$$\begin{pmatrix} \mathbf{A}_0 + \sum \lambda_i \mathbf{A}_i & \mathbf{B}_0 + \sum \lambda_i \mathbf{B}_i \\ (\mathbf{B}_0 + \sum \lambda_i \mathbf{B}_i)^T & \frac{1}{r}(c_0 + \sum \lambda_i c_i + t)\mathbf{I}_r + \mathbf{Z} \end{pmatrix} \succeq \mathbf{0},$$

and $\text{Tr}(\mathbf{Z}) \leq 0$. Making the change of variables $\Phi = \mathbf{Z} + \frac{t}{r}\mathbf{I}_r$, we deduce that the LMI (29) is equivalent to the existence of a matrix $\Phi \in \mathcal{S}^r$ such that

$$(30) \quad \begin{pmatrix} \mathbf{A}_0 + \sum \lambda_i \mathbf{A}_i & \mathbf{B}_0 + \sum \lambda_i \mathbf{B}_i \\ (\mathbf{B}_0 + \sum \lambda_i \mathbf{B}_i)^T & \frac{1}{r}(c_0 + \sum \lambda_i c_i)\mathbf{I}_r + \Phi \end{pmatrix} \succeq \mathbf{0},$$

and

$$(31) \quad \text{Tr}(\Phi) \leq t.$$

Replacing the LMI in problem (24) with the LMIs (30) and (31), problem (DV) is transformed into

$$\begin{aligned} & \max_{\lambda_i, \Phi, t} - \sum_{i \in \mathcal{I} \cup \mathcal{E}} \lambda_i \alpha_i - t \\ & \text{s.t.} \quad \begin{pmatrix} \mathbf{A}_0 + \sum \lambda_i \mathbf{A}_i & \mathbf{B}_0 + \sum \lambda_i \mathbf{B}_i \\ (\mathbf{B}_0 + \sum \lambda_i \mathbf{B}_i)^T & \frac{1}{r}(c_0 + \sum \lambda_i c_i)\mathbf{I}_r + \Phi \end{pmatrix} \succeq \mathbf{0}, \\ & \quad \lambda_i \geq 0, i \in \mathcal{I}, \\ & \quad \text{Tr}(\Phi) \leq t. \end{aligned}$$

It is clear that any optimal solution of the last problem satisfies $t = \text{Tr}(\Phi)$, and thus the problem is the same as

$$\begin{aligned} & \max_{\lambda_i, \Phi} - \sum_{i \in \mathcal{I} \cup \mathcal{E}} \lambda_i \alpha_i - \text{Tr}(\Phi) \\ & \text{s.t.} \begin{pmatrix} \mathbf{A}_0 + \sum \lambda_i \mathbf{A}_i & \mathbf{B}_0 + \sum \lambda_i \mathbf{B}_i \\ (\mathbf{B}_0 + \sum \lambda_i \mathbf{B}_i)^T & \frac{1}{r}(c_0 + \sum \lambda_i c_i) \mathbf{I}_r + \Phi \end{pmatrix} \succeq \mathbf{0}, \\ & \lambda_i \geq 0, i \in \mathcal{I}, \end{aligned}$$

which is identical to problem (DM). \square

Combining the latter result with the strong duality result, Corollary 3.6, the following corollary immediately follows.

COROLLARY 4.4. *Consider the QMP problem (6) with $|\mathcal{I}| + |\mathcal{E}| \leq r$, its vectorized semidefinite relaxation (SDRV) (problem (23)), and its vectorized dual (DV) (problem (24)). Suppose that condition (20) holds true and that the QMP problem is feasible. Then problems (QMP) and (SDRV) are solvable and $\text{val}(\text{QMP}) = \text{val}(\text{SDRV}) = \text{val}(\text{DV})$.*

5. An application to robust least squares. We continue the example from section 2.2. Suppose that the uncertainty set \mathcal{U} associated with the matrix \mathbf{A} is given by multiple norm constraints:

$$(32) \quad \mathcal{U} = \{ \Delta \in \mathbb{R}^{n \times r} : \|\mathbf{L}_i \Delta\|^2 \leq \rho_i, i = 1, \dots, m \},$$

where $\mathbf{L}_i \in \mathbb{R}^{k_i \times n}$ for some positive integers k_1, \dots, k_m and $\rho_i > 0, i = 1, \dots, m$. The above form of the uncertainty set is more general than the standard single-constraint form, and it can thus be used to describe more complicated scenarios of uncertainties. For example, by setting $k_i = n, m = n$, and $\mathbf{L}_i = \mathbf{E}_{ii}^n$, we model the situation in which the uncertainty associated with each column of the matrix \mathbf{A} has a separate norm constraint.

Assume that there exist nonnegative numbers $\gamma_1, \dots, \gamma_m$ such that

$$\sum_{i=1}^m \gamma_i \mathbf{L}_i^T \mathbf{L}_i \succ \mathbf{0}.$$

If $m \leq r$, then the conditions of Corollary 4.4 are satisfied, and as a consequence the inner maximization problem (9) is equal to the value of the dual problem given by

$$\begin{aligned} & \min_{t, \lambda_i} \sum_{i=1}^m \lambda_i \rho_i + t \\ & \text{s.t.} \begin{pmatrix} \mathbf{I}_r \otimes (-\mathbf{Q} + \sum_{i=1}^m \lambda_i \mathbf{L}_i^T \mathbf{L}_i) & -\text{vec}(\mathbf{F}) \\ -\text{vec}(\mathbf{F})^T & -c + t \end{pmatrix} \succeq \mathbf{0}, \\ & \lambda_i \geq 0, i = 1, 2, \dots, m. \end{aligned}$$

Here we considered the equivalent vectorized dual because it is not clear how to derive an SDP formulation from the nonvectorized dual. Now, using the identities (see [10])

$$\begin{aligned} \mathbf{I}_r \otimes \mathbf{Q} & \stackrel{(10)}{=} \mathbf{I}_r \otimes \mathbf{x} \mathbf{x}^T = (\mathbf{I}_r \otimes \mathbf{x})(\mathbf{I}_r \otimes \mathbf{x})^T, \\ \text{vec}(\mathbf{F}) & \stackrel{(10)}{=} \text{vec}(\mathbf{x}(\mathbf{A} \mathbf{x} - \mathbf{b})^T) = (\mathbf{I}_r \otimes \mathbf{x})(\mathbf{A} \mathbf{x} - \mathbf{b}), \end{aligned}$$

the dual problem is transformed into

$$\begin{aligned} & \min_{t, \lambda_i} \sum_{i=1}^m \lambda_i \rho_i + t \\ & \text{s.t.} \begin{pmatrix} -(\mathbf{I}_r \otimes \mathbf{x})(\mathbf{I}_r \otimes \mathbf{x})^T + \sum_{i=1}^m \lambda_i (\mathbf{I}_r \otimes (\mathbf{L}_i^T \mathbf{L}_i)) & -(\mathbf{I}_r \otimes \mathbf{x})(\mathbf{A}\mathbf{x} - \mathbf{b}) \\ -(\mathbf{A}\mathbf{x} - \mathbf{b})^T (\mathbf{I}_r \otimes \mathbf{x})^T & -\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + t \end{pmatrix} \succeq \mathbf{0}, \\ & \lambda_i \geq 0, i = 1, 2, \dots, m, \end{aligned}$$

which, by the Schur complement can be written as

$$\begin{aligned} & \min_{t, \lambda_i} \sum_{i=1}^m \lambda_i \rho_i + t \\ & \text{s.t.} \begin{pmatrix} \mathbf{I}_r & (\mathbf{I}_r \otimes \mathbf{x})^T & \mathbf{A}\mathbf{x} - \mathbf{b} \\ \mathbf{I}_r \otimes \mathbf{x} & \sum_{i=1}^m \lambda_i (\mathbf{I}_r \otimes (\mathbf{L}_i^T \mathbf{L}_i)) & \mathbf{0} \\ (\mathbf{A}\mathbf{x} - \mathbf{b})^T & \mathbf{0} & t \end{pmatrix} \succeq \mathbf{0}, \\ & \lambda_i \geq 0, i = 1, 2, \dots, m. \end{aligned}$$

Finally, we arrive at the following SDP formulation of the RLS problem (8):

$$\begin{aligned} & \min_{t, \lambda_i, \mathbf{x}} \sum_{i=1}^m \lambda_i \rho_i + t \\ & \text{s.t.} \begin{pmatrix} \mathbf{I}_r & (\mathbf{I}_r \otimes \mathbf{x})^T & \mathbf{A}\mathbf{x} - \mathbf{b} \\ \mathbf{I}_r \otimes \mathbf{x} & \sum_{i=1}^m \lambda_i (\mathbf{I}_r \otimes (\mathbf{L}_i^T \mathbf{L}_i)) & \mathbf{0} \\ (\mathbf{A}\mathbf{x} - \mathbf{b})^T & \mathbf{0} & t \end{pmatrix} \succeq \mathbf{0}, \\ & \lambda_i \geq 0, i = 1, 2, \dots, m. \end{aligned}$$

Appendix. A rank reduction algorithm for solvable semidefinite problems. We review here the rank reduction algorithm of [15] for solving SDP problems of the form (22).⁵ The underlying assumption that guarantees the validity of the process is that problem (22) is solvable and that $|\mathcal{I}_1| + |\mathcal{E}_1| \leq \binom{r+2}{2} - 1$.

ALGORITHM RED.

Input: \mathbf{X}_0 , an optimal solution to problem (22).

Output: An optimal solution \mathbf{X}^* to problem (22) satisfying $\text{rank}(\mathbf{X}^*) \leq r$.

1. **If** $\text{rank}(\mathbf{X}_0) \leq r$, then go to step 3. **Else** go to step 2.
2. **While** $\text{rank}(\mathbf{X}_0) > r$, **repeat** steps (a)–(e):
 - (a) **Set** $d \leftarrow \text{rank}(\mathbf{X}_0)$.
 - (b) Compute a decomposition of \mathbf{X}_0 : $\mathbf{X}_0 = \mathbf{U}\mathbf{U}^T$, where $\mathbf{U} \in \mathbb{R}^{n \times d}$.
 - (c) Find a nontrivial solution⁶ \mathbf{Z}_0 for the set of homogenous linear equations in the $d \times d$ symmetric variables matrix \mathbf{Z} ($\mathbf{Z} = \mathbf{Z}^T$):

$$\text{Tr}(\mathbf{U}^T \mathbf{C}_i \mathbf{U} \mathbf{Z}) = 0, \quad i \in \mathcal{I}_1 \cup \mathcal{E}_1.$$

- (d) **If** $\mathbf{Z}_0 \succeq \mathbf{0}$, then **set** $\mathbf{W} \leftarrow -\mathbf{Z}_0$. **Else set** $\mathbf{W} \leftarrow \mathbf{Z}_0$.
- (e) **Set** $\mathbf{X}_0 \leftarrow \mathbf{U}(\mathbf{I} + \beta \mathbf{W})\mathbf{U}^T$, where $\beta = -1/\lambda_{\min}(\mathbf{W})$.
3. **Set** $\mathbf{X}^* \leftarrow \mathbf{X}_0$ and **STOP**.

⁵Note that in [15], the SDP problem contains only inequality constraints. However, it is immediately seen that exactly the same rank reduction algorithm also works here.

⁶Using the relations $|\mathcal{I}_1| + |\mathcal{E}_1| \leq \binom{r+2}{2} - 1, d > r$, it is easy to see that the homogenous system has more variables than equations and, as a result, has a nonzero solution.

Note that the algorithm does not make use of the matrix C_0 corresponding to the objective function in (22). Indeed, it can be shown that since the input to the algorithm is an *optimal* solution of the SDP problem (22), then the value $\text{Tr}(C_0 X_0)$ remains constant throughout the process.

REFERENCES

- [1] K. ANSTREICHER AND H. WOLKOWICZ, *On Lagrangian relaxation of quadratic matrix constraints*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 41–55.
- [2] A. BARVINOK, *A remark on the rank of positive semidefinite matrices subject to affine constraints*, Discrete Comput. Geom., 25 (2001), pp. 23–31.
- [3] A. BARVINOK, *Problems of distance geometry and convex properties of quadratic maps*, Discrete Comput. Geom., 13 (1995), pp. 189–202.
- [4] A. BEN-TAL AND A. NEMIROVSKI, *Lectures on Modern Convex Optimization, Analysis, Algorithms, and Engineering Applications*, MPS-SIAM Ser. Optim. 2, SIAM, Philadelphia, 2001.
- [5] A. BEN-TAL AND M. TEBoulLE, *Hidden convexity in some nonconvex quadratically constrained quadratic programming*, Math. Programming, 72 (1996), pp. 51–63.
- [6] S. CHANDRASEKARAN, G. H. GOLUB, M. GU, AND A. H. SAYED, *Parameter estimation in the presence of bounded data uncertainties*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 235–252.
- [7] L. EL DÉN AND H. PARK, *A Procrustes problem on the Stiefel manifold*, Numer. Math., 82 (1999), pp. 599–619.
- [8] C. FORTIN AND H. WOLKOWICZ, *The trust region subproblem and semidefinite programming*, Optim. Methods Softw., 19 (2004), pp. 41–67.
- [9] L. EL GHAOU AND H. LEBRET, *Robust solutions to least-squares problems with uncertain data*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 1035–1064.
- [10] A. GRAHAM, *Kronecker Products and Matrix Calculus: With Applications*, Ellis Horwood Ser. Math. Appl., Ellis Horwood, Chichester, U.K., 1981.
- [11] J. B. HIRIART-URRUTY AND M. TORKI, *Permanently going back and forth between the “quadratic world” and the “convexity world” in optimization*, Appl. Math. Optim., 45 (2002), pp. 169–184.
- [12] J. J. MORÉ, *Generalization of the trust region problem*, Optim. Methods Softw., 2 (1993), pp. 189–209.
- [13] J. J. MORÉ AND D. C. SORENSEN, *Computing a trust region step*, SIAM J. Sci. Statist. Comput., 4 (1983), pp. 553–572.
- [14] G. PATAKI, *On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues*, Math. Oper. Res., 23 (1998), pp. 339–358.
- [15] G. PATAKI, *The geometry of semidefinite programming*, in Handbook of Semidefinite Programming, Internat. Ser. Oper. Res. Management Sci. 27, Kluwer Academic Publishers, Boston, 2000, pp. 29–65.
- [16] B. T. POLYAK, *Convexity of quadratic transformations and its use in control and optimization*, J. Optim. Theory Appl., 99 (1998), pp. 553–583.
- [17] P. H. SCHÖNEMANN, *A generalized solution of the orthogonal Procrustes problem*, Psychometrika, 31 (1966), pp. 1–10.
- [18] R. J. STERN AND H. WOLKOWICZ, *Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations*, SIAM J. Optim., 5 (1995), pp. 286–313.
- [19] Y. YE AND S. ZHANG, *New results on quadratic minimization*, SIAM J. Optim., 14 (2003), pp. 245–267.
- [20] Y. YUAN, *On a subproblem of trust region algorithms for constrained optimization*, Math. Programming, 47 (1990), pp. 53–63.