A convex optimization approach for minimizing the ratio of indefinite quadratic functions over an ellipsoid

Amir Beck · **Marc Teboulle**

Abstract   We consider the nonconvex problem (RQ) of minimizing the ratio of two nonconvex quadratic functions over a possibly degenerate ellipsoid. This formulation is motivated by the so-called regularized total least squares problem (RTLS), which is a special case of the problem's class we study. We prove that under a certain mild assumption on the problem's data, problem (RQ) admits an exact semidefinite programming relaxation. We then study a simple iterative procedure which is proven to converge superlinearly to a global solution of (RQ) and show that the dependency of the number of iterations on the optimality tolerance ε grows as $O(\sqrt{\ln \varepsilon^{-1}})$.

Keywords   Ratio of quadratic minimization · Nonconvex quadratic minimization · Semidefinite programming · Strong duality · Regularized total least squares · Fixed point algorithms · Convergence analysis

Mathematics Subject Classification (2000)   90C22 · 90C25 · 62G05

A. Beck (✉)
Department of Industrial Engineering, Technion, Israel Institute of Technology,
Haifa 32000, Israel
e-mail: becka@ie.technion.ac.il

M. Teboulle
School of Mathematical Sciences, Tel-Aviv University,
Ramat-Aviv, Tel-Aviv 69978, Israel
e-mail: teboulle@post.tau.ac.il

1 Introduction

Convex optimization problems play a fundamental role in the theory and practice of continuous optimization. One major attribute associated with convex optimization problems is that they are computationally tractable for a wide class of problems. In particular, very efficient e.g., polynomial algorithms, are available to solve specific classes of convex problems such as linear programming, conic quadratic, semidefinite programming and more, see [16], and the more recent book [7].

In sharp contrast to the convex optimization setting, no efficient universal solution methods for solving nonconvex optimization problems are known, and there are strong reasons to believe that no such methods exist. Nonetheless, there are some classes of nonconvex problems that possess a *hidden convexity* property, namely, they can be shown to have either a zero duality gap or to be equivalent to some convex optimization reformulation, and as such are tractable. The simplest and well known example is the trust region problem, which consists of minimizing an indefinite quadratic function over a ball, and which admits an exact semidefinite reformulation, see e.g., the work [10] and references therein. Extensions of this problem were considered in [15] and [22], where necessary and sufficient global optimality conditions for the problem of minimizing an indefinite quadratic function subject to a single indefinite homogeneous quadratic constraint, or two-sided homogeneous quadratic forms constraints were derived, which lead them to establish a duality result with no gap, and thus revealing the implicit convex nature of such problems. Moreover, in [8] it was proven that this problem admits a convex equivalent formulation via a simple a transformation, thus showing that it enjoys such a hidden convexity property. More recent results identifying further interesting classes of quadratic problems whose semidefinite relaxations admit no gap with the true optimal value can be found for example in [4,6,17,23] and references therein, as well as in [1,2] in the context of problems with quadratic matrix constraints.

In this paper we study a different class of nonconvex optimization problems which possesses a hidden convexity property, i.e., which will be shown to be equivalent to a convex problem and thus can be efficiently solved. More precisely, we study nonconvex problems involving an objective function which is a ratio of two quadratic functions and a single homogeneous quadratic constraint (a possibly degenerate ellipsoid):

$$\text{(RQ)} \quad \inf \left\{ \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} : \quad \|\mathbf{L}\mathbf{x}\|^2 \le \rho, \mathbf{x} \in \mathbb{R}^n \right\},$$

where f_i are nonconvex quadratic functions, ρ is a given positive scalar and $\mathbf{L}$ is a given $r \times n$ matrix, see Sect.1 for a precise formulation. The special case $f_2(\mathbf{x}) \equiv 1$ shows that the class of problems (RQ) includes problems of the form

$$\text{(GTRS)} \quad \inf \left\{ f_1(\mathbf{x}) : \quad \|\mathbf{L}\mathbf{x}\|^2 \le \rho, \mathbf{x} \in \mathbb{R}^n \right\}, \tag{1.1}$$

which are generalized trust region subproblems (GTRS) (i.e., not with a simple Euclidean norm constraint), see [15].

1.1 Motivation

A very interesting motivation for considering the class of Problems (RQ) is the regularized total least squares problem which we now briefly recall . The *total least squares* (TLS) [12,14] is a technique for dealing with linear system $\mathbf{Ax} \approx \mathbf{b}$, where both the matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and the vector $\mathbf{b} \in \mathbb{R}^m$ are contaminated by noise. The TLS approach to this problem is to seek a perturbation matrix $\mathbf{E} \in \mathbb{R}^{m \times n}$ and a perturbation vector $\mathbf{r} \in \mathbb{R}^m$ that minimize $\|\mathbf{E}\|^2 + \|\mathbf{r}\|^2$ subject to the consistency equation $(\mathbf{A} + \mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{r}$. The TLS approach was extensively used in a variety of scientific disciplines such as signal processing, automatic control, statistics, physics, economic, biology and medicine (see e.g., [14] and references therein). The TLS problem can be recast as an unconstrained quadratic fractional problem: [5,12,14]:

$$(TLS) : \quad \inf_{\mathbf{x} \in \mathbb{R}^n} \frac{\|\mathbf{Ax} - \mathbf{b}\|^2}{\|\mathbf{x}\|^2 + 1}.$$

In the case of ill-posed problems, the TLS approach might produce a solution with poor quality and thus regularization is introduced in order to stabilize the solution. A well-studied approach for this problem is the regularized TLS problem (RTLS) [5,11,13,18,20] in which a quadratic constraint is added in order to bound the size of the constraint:

$$(RTLS) : \quad \inf_{\mathbf{x} \in \mathbb{R}^n} \left\{ \frac{\|\mathbf{Ax} - \mathbf{b}\|^2}{\|\mathbf{x}\|^2 + 1} : \|\mathbf{Lx}\|^2 \leq \rho \right\}.$$

Currently, there exist several methods to tackle the RTLS problem [5,11,13,18,20]. Among them, the recent approach proposed in [5] appears to be the only one proven to converge to a global minimum. The procedure devised in [5] relies on a combination of a key observation due to [9] for fractional programs, and of the hidden convexity result of [8] just alluded above, and involves the solution of a GTRS problem at each iteration. As shown in [5], it leads to a practically efficient method for solving RTLS. In [20], a different iterative scheme was proposed, which also involves the solution of a GTRS problem. The numerical results presented in [20] indicate that the method converges at a very fast rate and requires the solution of very few (up to 5) GTRS problems. The numerical results reported in [20] also indicate that the method produces a *global solution*. This fact was also validated empirically by comparing the two procedures in [5]. However, a proof of convergence to a global optimal solution of the RTLS was not given in [20].

1.2 Main contributions

The aforementioned results suggest that the problem (RQ) of minimizing a quadratically constrained ratio of two quadratic functions seems to share some kind of hidden convexity property. In this paper we show that this is indeed the case. In the next section, we derive a simple condition in terms of the problem's data under which the

attainment of the minimum in problem (RQ) is warranted. This condition allows us to derive an appropriate nonconvex reformulation of (RQ), and to apply a strong duality result for nonconvex homogeneous quadratic problems, proven in [17]. By so doing, we prove in Sect. 3 that problem (RQ) can be recast as a semidefinite programming problem with no gap, and that the global optimal solution of the original problem (RQ) can be extracted from the optimal solution of this semidefinite formulation. In Sect. 4 we then study and extend the iterative scheme suggested in [20] for the RTLS, to the more general class of problems (RQ), and we develop a global convergence analysis. More precisely, we prove that this algorithm is superlinearly convergent. Moreover, we show that it produces an ε-global optimal solution to (RQ) after at most $A + B\sqrt{\ln \varepsilon^{-1}}$ iterations, for some positive constants A, B. These results also provide a theoretical justification of the successful computational results reported in the context of (RTLS) in [20].

1.3 Notation

Vectors are denoted by boldface lowercase letters, e.g., $\mathbf{y}$, and matrices are denoted by boldface uppercase letters *e.g.,* $\mathbf{A}$. For any symmetric matrix $\mathbf{A}$ and symmetric positive definite matrix $\mathbf{B}$ we denote the corresponding minimum generalized eigenvalue by $\lambda_{\min}(\mathbf{A}, \mathbf{B})$; the minimum generalized eigenvalue has several equivalent formulations:

$$\lambda_{\min}(\mathbf{A}, \mathbf{B}) = \max\{\lambda : \mathbf{A} - \lambda\mathbf{B} \succeq \mathbf{0}\} = \min_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^T \mathbf{A}\mathbf{x}}{\mathbf{x}^T \mathbf{B}\mathbf{x}} = \lambda_{\min}(\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2}),$$

where we use the notation $\mathbf{A} \succeq 0$ ($\mathbf{A} \succ 0$) for a positive semidefinite (positive definite) matrix $\mathbf{A}$. We also use some standard abbreviations such as SDP (semidefinite programming) and LMI (linear matrix inequalities). We also follow the MATLAB convention and use ";" for adjoining scalars, vectors or matrices in a column. Finally, we will make use of some standard convex analysis notation and definition. For an extended real-valued function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $\text{dom } f = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) < \infty\}$ denotes its effective domain, and epi $f = \{(\mathbf{x}, \mu) \in \mathbb{R}^n \times \mathbb{R} \mid \mu \geq f(\mathbf{x})\}$ is the epigraph of f. Recall that f is called proper whenever $\text{dom } f \neq \emptyset$ and there exists $\mathbf{y}$ such that $f(\mathbf{y}) > -\infty$, and f is lower semi-continuous (lsc) on $\mathbb{R}^n$ whenever $\liminf_{\mathbf{x} \to \mathbf{z}} f(\mathbf{x}) = f(\mathbf{z})$, for all $\mathbf{z} \in \mathbb{R}^n$.

2 Existence of a global minimizer

2.1 Problem formulation and basic properties

Consider the problem of minimizing the ratio of two quadratic functions subject to a single convex homogeneous constraint:

$$\text{(RQ)} \quad f^* = \inf_{\mathbf{x} \in \mathbb{R}^n} \left\{ f(\mathbf{x}) \equiv \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} : \|\mathbf{L}\mathbf{x}\|^2 \leq \rho \right\}, \tag{2.1}$$

where

$$f_i(\mathbf{x}) = \mathbf{x}^T \mathbf{A}_i \mathbf{x} + 2\mathbf{b}_i^T \mathbf{x} + c_i, \quad i = 1, 2,$$

with $\mathbf{A}_i$ a symmetric matrix, $\mathbf{b}_i \in \mathbb{R}^n$ and $c_i \in \mathbb{R}$, $i = 1, 2$. Moreover, $\mathbf{L}$ is an $r \times n$ ($r \leq n$) full row rank matrix and ρ is a positive number. Problem (RQ) is well-defined if and only if f_2 is not equal to zero over the feasible set. Therefore, the following assumption is made throughout the paper:

Assumption 1 There exists $\eta \geq 0$ such that

$$\begin{pmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^T & c_2 \end{pmatrix} + \eta \begin{pmatrix} \mathbf{L}^T\mathbf{L} & \mathbf{0} \\ \mathbf{0} & -\rho \end{pmatrix} \succ \mathbf{0}. \tag{2.2}$$

It is easy to see that Assumption 1 implies that $f_2(\mathbf{x}) > 0$ for any feasible point $\mathbf{x}$, (see Proposition 1(iii)) so that problem (RQ) is well-defined. Moreover, this assumption is satisfied for the GTRS problem when $r = n$ (for $0 < \eta < 1/\rho$) and for the RTLS problem in which $f_2(\mathbf{x}) = \|\mathbf{x}\|^2 + 1$.

Several useful consequences can be drawn from the LMI (2.2) which we collect in the next result.

Proposition 1 *Under Assumption* 1 *the following statements hold*:

  (i) $c_2 > 0$.
 (ii) *Let* $\mathbf{F} \in \mathbb{R}^{n \times (n-r)}$ *be a matrix whose columns form an orthonormal basis for the null space of* $\mathbf{L}$. *Then*

$$\begin{pmatrix} \mathbf{F}^T \mathbf{A}_2 \mathbf{F} & \mathbf{F}^T \mathbf{b}_2 \\ \mathbf{b}_2^T \mathbf{F} & c_2 \end{pmatrix} \succ \mathbf{0},$$

     and in particular $\mathbf{F}^T \mathbf{A}_2 \mathbf{F} \succ \mathbf{0}$.
(iii)

$$f_2(\mathbf{x}) \geq \delta(\|\mathbf{x}\|^2 + 1) \geq \delta > 0 \text{ for every } \mathbf{x} \text{ such that } \|\mathbf{L}\mathbf{x}\|^2 \leq \rho, \tag{2.3}$$

    where

$$\delta = \lambda_{\min}\left[\begin{pmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^T & c_2 \end{pmatrix} + \eta \begin{pmatrix} \mathbf{L}^T\mathbf{L} & \mathbf{0} \\ \mathbf{0} & -\rho \end{pmatrix} \right]. \tag{2.4}$$

Proof (i) Since the matrix in the LHS of (2.2) is positive definite it follows that all of its diagonal elements are positive and in particular the $(n + 1, n + 1)$th entry is positive: $c_2 - \eta\rho > 0$, which implies that $c_2 > 0$.
(ii) By multiplying the LMI (2.2) from the left by $\mathbf{G}^T$ and from the right by $\mathbf{G}$ where

$$\mathbf{G} = \begin{pmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix},$$

we obtain (recall that if $\mathbf{M} \succ \mathbf{0}$ then $\mathbf{G}^T \mathbf{M} \mathbf{G} \succ \mathbf{0}$ if $\mathbf{G}$ has full row rank)

$$\begin{pmatrix} \mathbf{F}^T \mathbf{A}_2 \mathbf{F} & \mathbf{F}^T \mathbf{b}_2 \\ \mathbf{b}_2^T \mathbf{F} & c_2 \end{pmatrix} + \eta \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\rho \end{pmatrix} \succ \mathbf{0}.$$

Since $\rho > 0$ and $\eta \geq 0$ we conclude that

$$\begin{pmatrix} \mathbf{F}^T \mathbf{A}_2 \mathbf{F} & \mathbf{F}^T \mathbf{b}_2 \\ \mathbf{b}_2^T \mathbf{F} & c_2 \end{pmatrix} \succ \mathbf{0},$$

which also implies that $\mathbf{F}^T \mathbf{A}_2 \mathbf{F} \succ \mathbf{0}$.

(iii) It follows immediately from (2.2) that

$$\begin{pmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^T & c_2 \end{pmatrix} + \eta \begin{pmatrix} \mathbf{L}^T \mathbf{L} & \mathbf{0} \\ \mathbf{0} & -\rho \end{pmatrix} \succeq \delta \mathbf{I} \qquad (2.5)$$

with δ given in (2.4). Multiplying (2.5) by $(\mathbf{x}^T, 1)$ from the left, and by $(\mathbf{x}^T, 1)^T$ from the right results in

$$f_2(\mathbf{x}) + \eta(\|\mathbf{L}\mathbf{x}\|^2 - \rho) \geq \delta(\|\mathbf{x}\|^2 + 1),$$

which readily implies (2.3). □

We now show that Assumption 1 guarantees the finiteness of the infimum in problem (RQ) defined in (2.1).

Lemma 1 *The infimum of problem* (2.1) *is finite.*

Proof Define,

$$d_1 = \inf\{f(\mathbf{x}) : \|\mathbf{L}\mathbf{x}\|^2 \leq \rho, \, f_1(\mathbf{x}) \geq 0\},$$
$$d_2 = \inf\{f(\mathbf{x}) : \|\mathbf{L}\mathbf{x}\|^2 \leq \rho, \, f_1(\mathbf{x}) \leq 0\}.$$

Using the simple relation

$$\inf\{f(x) : x \in C_1 \cup C_2\} = \min\{\inf_{x \in C_1} f(x), \inf_{x \in C_2} f(x)\},$$

one has $f^* = \min\{d_1, d_2\}$. By its definition, d_1 is nonnegative. It remains to show that d_2 is finite. Indeed, for every $\mathbf{x}$ satisfying $\|\mathbf{L}\mathbf{x}\|^2 \leq \rho$ and $f_1(\mathbf{x}) \leq 0$, we have

$$f(\mathbf{x}) = \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} \geq \frac{f_1(\mathbf{x})}{\delta(\|\mathbf{x}\|^2 + 1)} \geq \frac{1}{\delta} \lambda_{\min} \begin{pmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^T & c_1 \end{pmatrix} \equiv l,$$

where the first inequality is due to $f_1(\mathbf{x}) \leq 0$ and (2.3). Therefore, $d_2 \geq l$, and hence $f^* \geq \min\{0, l\}$ is finite. □

Note that in the RQ problem we make no assumptions on the convexity of either the numerator function f_1 or the denominator function f_2 (as opposed to the RTLS problem), and the analysis to follow does not rely on any convexity assumptions on the terms defining the objective function of (RQ).

2.2 Attainment of the minimum

Attainment of the minimum of problem (RQ) is not always guaranteed. When the matrix $\mathbf{L}$ is a non-singular square matrix ($r = n$), then the feasible region is a nondegenerate ellipsoid (i.e., a compact set) and hence the minimum is attained. However, in many interesting applications, such as in RTLS problems, one has $r < n$. For example, this occurs when $\mathbf{L}$ represents discretization of first or second order differential operators, see [11,20].

Example 1 Consider problem (RQ) with

$$\mathbf{A}_1 = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 1 \end{pmatrix}, \ \mathbf{b}_1 = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \ c_1 = 5,$$

$$\mathbf{A}_2 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}, \ \mathbf{b}_2 = \mathbf{0}, \qquad c_2 = 1,$$

$$\mathbf{L} = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad \rho = 1.$$

In this case, problem (RQ) has the following form:

$$\inf_{x_1,x_2} \left\{ f(x_1, x_2) = \frac{5 - 4x_1 + 2x_1^2 + x_2^2 + x_1 x_2}{1 + x_1^2 + x_2^2 + x_1 x_2} : x_1^2 \le 1 \right\}. \tag{2.6}$$

To show that the minimum of problem (2.6) is not attained, note that for every x_1 such that $x_1^2 \le 1$ we have

$$f(x_1, x_2) = 1 + \frac{(x_1 - 2)^2}{1 + x_1^2 + x_2^2 + x_1 x_2} > 1.$$

On the other hand, the infimum is equal to 1 since $f(0, x_2) \to 1$ as x_2 tends to ∞. Note that the minimum of the *unconstrained* problem is attained; the minimum is equal to 1 and attained at any point of the form $(2, x_2)$.

We will prove that attainment of the minimum in the case where $r < n$ can be guaranteed under a certain mild sufficient condition. For that purpose we first recall some useful preliminary results on asymptotic cones and functions, see e.g., [3] for details and proofs.

Definition 1 Let C be a nonempty set in $\mathbb{R}^n$. The asymptotic cone of the set C, denoted by C_∞, is defined by:

$$C_\infty = \left\{ \mathbf{d} \in \mathbb{R}^n \mid \exists t_k \to +\infty, \exists \mathbf{x}_k \in C \text{ with } \lim_{k \to \infty} \frac{\mathbf{x}_k}{t_k} = \mathbf{d} \right\}.$$

As an immediate useful consequence, a set $C \subseteq \mathbb{R}^n$ is bounded if and only if $C_\infty = \{\mathbf{0}\}$.

We also need the concept of asymptotic function for an arbitrary *nonconvex* function. Recall that for any proper function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, there exists a unique

function $f_\infty : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ associated with f, called the asymptotic function, such that epi $f_\infty = (\text{epi } f)_\infty$. A fundamental analytic representation of the asymptotic function f_∞ is given by (see, [3, Theorem 2.5.1, p. 49])

$$f_\infty(\mathbf{d}) = \liminf_{\substack{\mathbf{d}' \to \mathbf{d} \\ t \to +\infty}} t^{-1} f(t\mathbf{d}'). \tag{2.7}$$

Note that when f is assumed proper lsc and convex, the above formula simplifies to

$$f_\infty(\mathbf{d}) = \lim_{t \to +\infty} t^{-1} f(t\mathbf{d}), \ \forall \mathbf{d} \in \text{dom } f, \tag{2.8}$$

and if $\mathbf{0} \in \text{dom } f$, the formula holds for every $\mathbf{d} \in \mathbb{R}^n$.

Using the above concepts and results, we obtain,

Lemma 2 *Let* $\alpha, \delta \in \mathbb{R}$ *and define*

$$C := \{\mathbf{x} \in \mathbb{R}^n \mid f_1(\mathbf{x}) - \alpha f_2(\mathbf{x}) \leq \delta, \|\mathbf{L}\mathbf{x}\|^2 \leq \rho\}.$$

Then,

(i) $C_\infty \subseteq \{\mathbf{d} \in \mathbb{R}^n \mid \mathbf{d}^T (\mathbf{A}_1 - \alpha \mathbf{A}_2)\mathbf{d} \leq 0\} \cap \text{Ker}(\mathbf{L}).$

(ii) *Let* $\mathbf{F} \in \mathbb{R}^{n \times (n-r)}$ *be a matrix whose columns form an orthonormal basis for the null space of* $\mathbf{L}$. *If* $\alpha < \lambda_{\min}(\mathbf{F}^T \mathbf{A}_1 \mathbf{F}, \mathbf{F}^T \mathbf{A}_2 \mathbf{F})$ *then* C *is a compact set.*

Proof (i) Using the definition of f_1 and f_2, one has

$$C = \{\mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) \leq 0, \ q(\mathbf{x}) \leq 0\},$$

where

$$g(\mathbf{x}) := \mathbf{x}^T (\mathbf{A}_1 - \alpha \mathbf{A}_2)\mathbf{x} + 2(\mathbf{b}_1 - \alpha \mathbf{b}_2)^T \mathbf{x} + c_1 - \alpha c_2 - \delta, \ q(\mathbf{x}) = \mathbf{x}^T \mathbf{L}^T \mathbf{L} \mathbf{x} - \rho.$$

Therefore, invoking [3, Corollary 2.5.4, p.52], one has

$$C_\infty \subseteq \{\mathbf{d} \in \mathbb{R}^n \mid g_\infty(\mathbf{d}) \leq 0, \ q_\infty(\mathbf{d}) \leq 0\}.$$

Using (2.8) and (2.7) it can be verified that (see e.g. [3, Example 2.5.1, p. 51]):

$$q_\infty(\mathbf{d}) = 2\mathbf{b}^T \mathbf{d} + \delta(\mathbf{d} | \text{Ker } \mathbf{L}^T \mathbf{L}) = 2\mathbf{b}^T \mathbf{d} + \delta(\mathbf{d} | \text{Ker } \mathbf{L}),$$

and

$$g_\infty(\mathbf{d}) = \begin{cases} -\infty & \text{if } \mathbf{d}^T (\mathbf{A}_1 - \alpha \mathbf{A}_2)\mathbf{d} \leq 0, \\ +\infty & \text{if } \mathbf{d}^T (\mathbf{A}_1 - \alpha \mathbf{A}_2)\mathbf{d} > 0. \end{cases}$$

from which the desired inclusion (i) follows.

(ii) Recall that $\mathbf{d} \in \text{Ker} \, \mathbf{L}$ if and only if $\mathbf{d} = \mathbf{F}\mathbf{v}$ for some $\mathbf{v} \in \mathbb{R}^{n-r}$. Let $\mathbf{d} \in C_\infty$ and suppose that $\mathbf{d} \neq \mathbf{0}$. Then, applying the first part of the lemma, it follows that

$$\alpha \geq \lambda_{\min}(\mathbf{F}^T \mathbf{A}_1 \mathbf{F}, \mathbf{F}^T \mathbf{A}_2 \mathbf{F}),$$

which contradicts our hypothesis $\alpha < \lambda_{\min}(\mathbf{F}^T \mathbf{A}_1 \mathbf{F}, \mathbf{F}^T \mathbf{A}_2 \mathbf{F})$. Therefore, one must have $\mathbf{d} = \mathbf{0}$, and hence $C_\infty \subseteq \{\mathbf{0}\}$, from which the compactness of C follows. $\square$

We are now ready to state and prove the main result of this section.

Theorem 1 *Consider the RQ problem* (2.1) *with $r < n$. Suppose that the following condition is satisfied*:

$$\lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) < \lambda_{\min}(\mathbf{F}^T \mathbf{A}_1 \mathbf{F}, \mathbf{F}^T \mathbf{A}_2 \mathbf{F}), \tag{2.9}$$

where

$$\mathbf{M}_1 = \begin{pmatrix} \mathbf{F}^T \mathbf{A}_1 \mathbf{F} & \mathbf{F}^T \mathbf{b}_1 \\ \mathbf{b}_1^T \mathbf{F} & c_1 \end{pmatrix}, \mathbf{M}_2 = \begin{pmatrix} \mathbf{F}^T \mathbf{A}_2 \mathbf{F} & \mathbf{F}^T \mathbf{b}_2 \\ \mathbf{b}_2^T \mathbf{F} & c_2 \end{pmatrix} \tag{2.10}$$

and $\mathbf{F}$ is an $n \times (n-r)$ matrix whose columns form an orthonormal basis for the null space of L. Then the following statements hold

(i) *Any minimum generalized eigenvector $(\bar{\mathbf{v}}; \bar{t})$ of the matrix pair $\mathbf{M}_1, \mathbf{M}_2$ ($\bar{\mathbf{v}} \in \mathbb{R}^{n-r}, \bar{t} \in \mathbb{R}$) satisfies $\bar{t} \neq 0$ and $f(\mathbf{F}\bar{\mathbf{v}}/\bar{t}) = \lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2)$.*

(ii)

$$f^* \leq \lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2). \tag{2.11}$$

(iii) *The minimum of problem (RQ) is attained.*

Proof (i) Using the definition of the minimum generalized eigenvalue we have

$$\lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) = \min_{\mathbf{v} \in \mathbb{R}^{n-r}, t \in \mathbb{R}} \left\{ \frac{\mathbf{v}^T \mathbf{F}^T \mathbf{A}_1 \mathbf{F}\mathbf{v} + 2\mathbf{b}_1^T \mathbf{F}\mathbf{v}t + c_1 t^2}{\mathbf{v}^T \mathbf{F}^T \mathbf{A}_2 \mathbf{F}\mathbf{v} + 2\mathbf{b}_2^T \mathbf{F}\mathbf{v}t + c_2 t^2} : (\mathbf{v}; t) \neq \mathbf{0} \right\}. \tag{2.12}$$

The minimum in the latter minimization problem is attained at $(\bar{\mathbf{v}}; \bar{t})$. In order to show that $\bar{t} \neq 0$ we assume by contradiction that the minimum in (2.12) is attained at $t = 0$. In that case we would have

$$\lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) = \min_{\mathbf{v} \in \mathbb{R}^{n-r}, \mathbf{v} \neq \mathbf{0}} \frac{\mathbf{v}^T \mathbf{F}^T \mathbf{A}_1 \mathbf{F}\mathbf{v}}{\mathbf{v}^T \mathbf{F}^T \mathbf{A}_2 \mathbf{F}\mathbf{v}} = \lambda_{\min}(\mathbf{F}^T \mathbf{A}_1 \mathbf{F}, \mathbf{F}^T \mathbf{A}_2 \mathbf{F}),$$

which contradicts condition (2.9). Finally,

$$\begin{aligned} f(\mathbf{F}\bar{\mathbf{v}}/\bar{t}) &= \frac{(\bar{\mathbf{v}}/\bar{t})^T \mathbf{F}^T \mathbf{A}_1 \mathbf{F}(\bar{\mathbf{v}}/\bar{t}) + 2\mathbf{b}_1^T \mathbf{F}(\bar{\mathbf{v}}/\bar{t}) + c_1}{(\bar{\mathbf{v}}/\bar{t})^T \mathbf{F}^T \mathbf{A}_2 \mathbf{F}(\bar{\mathbf{v}}/\bar{t}) + \mathbf{b}_2^T \mathbf{F}(\bar{\mathbf{v}}/\bar{t}) + c_2} \\ &= \frac{\bar{\mathbf{v}}^T \mathbf{F}^T \mathbf{A}_1 \mathbf{F}\bar{\mathbf{v}} + 2\mathbf{b}_1^T \mathbf{F}\bar{\mathbf{v}}\bar{t} + c_1 \bar{t}^2}{\bar{\mathbf{v}}^T \mathbf{F}^T \mathbf{A}_2 \mathbf{F}\bar{\mathbf{v}} + \mathbf{b}_2^T \mathbf{F}\bar{\mathbf{v}}\bar{t} + c_2 \bar{t}^2} = \lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2). \end{aligned}$$

(ii) Follows from the fact that $\mathbf{F}\bar{\mathbf{v}}/\bar{t} \in \mathrm{Ker}\,\mathbf{L}$ is a feasible point of (2.1).

(iii) Let $\mathbf{x}_0 = \mathbf{F}\bar{\mathbf{v}}/\bar{t}$ be a feasible point of problem (RQ), set $\alpha = f(\mathbf{x}_0)$ and define:

$$C := \{\mathbf{x} \in \mathbb{R}^n \mid f_1(\mathbf{x}) - \alpha f_2(\mathbf{x}) \le 0, \|\mathbf{L}\mathbf{x}\|^2 \le \rho\}.$$

Since C is the intersection of the feasible set of problem (RQ) with a nonempty level set of its objective function f, we conclude that problem (RQ) is equivalent to solving $\min\{f(\mathbf{x}) \mid \mathbf{x} \in C\}$, and hence it remains to show that C is bounded. Since under our assumption (2.9) one has

$$\alpha = f(\mathbf{x}_0) = \lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) < \lambda_{\min}(\mathbf{F}^T\mathbf{A}_1\mathbf{F}, \mathbf{F}^T\mathbf{A}_2\mathbf{F}),$$

it follows by the second part of Lemma 2 that C is compact. □

Remark 1 (i) *Weak* inequality is always satisfied in the condition (2.9) since

$$\lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) = \min_{(\mathbf{v};t)\neq 0} \frac{\mathbf{v}^T\mathbf{F}^T\mathbf{A}_1\mathbf{F}\mathbf{v} + 2\mathbf{b}_1^T\mathbf{F}\mathbf{v}t + c_1 t^2}{\mathbf{v}^T\mathbf{F}^T\mathbf{A}_2\mathbf{F}\mathbf{v} + 2\mathbf{b}_2^T\mathbf{F}\mathbf{v}t + c_2 t^2}$$

$$\le \min_{\mathbf{v}\neq 0} \frac{\mathbf{v}^T\mathbf{F}^T\mathbf{A}_1\mathbf{F}\mathbf{v}}{\mathbf{v}^T\mathbf{F}^T\mathbf{A}_2\mathbf{F}\mathbf{v}} = \lambda_{\min}(\mathbf{F}^T\mathbf{A}_1\mathbf{F}, \mathbf{F}^T\mathbf{A}_2\mathbf{F}).$$

(ii) A direct consequence of (2.9) and (2.11) is that

$$f^* < \lambda_{\min}(\mathbf{F}^T\mathbf{A}_1\mathbf{F}, \mathbf{F}^T\mathbf{A}_2\mathbf{F}). \tag{2.13}$$

(iii) For $\mathbf{L} = \mathbf{0}$ and $f_2(\mathbf{x}) = \|\mathbf{x}\|^2 + 1$, problem (2.1) reduces to the classical (unconstrained) TLS problem. In this case we can take $\mathbf{F} = \mathbf{I}$ in condition (2.9), which then reduces to the well known condition for the attainability of the minimum in the TLS problem:

$$\lambda_{\min}\begin{pmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{A}^T\mathbf{b} \\ \mathbf{b}^T\mathbf{A} & \|\mathbf{b}\|^2 \end{pmatrix} < \lambda_{\min}(\mathbf{A}^T\mathbf{A}). \tag{2.14}$$

In that case, any minimum eigenvector $(\mathbf{v}; t)$ of the matrix $\begin{pmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{A}^T\mathbf{b} \\ \mathbf{b}^T\mathbf{A} & \|\mathbf{b}\|^2 \end{pmatrix}$ satisfies $t \neq 0$ and $\mathbf{v}/t$ is a solution to the TLS problem [12,14].

Example 2 (**Continuation of Example 1**) Since the minimum in problem (2.6) is not attained, condition (2.9) does not hold true. To show this, note that here $\mathbf{F}$ can be chosen as $(0; 1)$ and thus

$$\mathbf{F}^T\mathbf{A}_1\mathbf{F} = 1, \mathbf{F}^T\mathbf{A}_2\mathbf{F} = 1, \mathbf{M}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}, \mathbf{M}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence,

$$\lambda_{\min}(\mathbf{F}^T\mathbf{A}_1\mathbf{F}, \mathbf{F}^T\mathbf{A}_2\mathbf{F}) = 1, \lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) = 1,$$

which shows that condition (2.9) is not satisfied.

We have shown in this section that if the following condition is satisfied:

[SC] : Either $(r = n)$ or $(r < n$ and $\lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) < \lambda_{\min}(\mathbf{F}^T \mathbf{A}_1 \mathbf{F}, \mathbf{F}^T \mathbf{A}_2 \mathbf{F}))$,

$$(2.15)$$

then the minimum of the RQ problem (2.1) is attained. This will be considered as a blanket hypothesis in the remaining parts of the paper.

3 Exact SDP relaxation for (RQ)

In this section we prove that under our blanket hypothesis [SC], the nonconvex problem (RQ) admits an exact convex semidefinite formulation.

3.1 Formulation of (RQ) as a non-convex quadratic programming problem

We first show that problem (RQ) is equivalent to solving a nonconvex homogeneous quadratic optimization problem.

Lemma 3 *Suppose that condition* [*SC*] *given in* (2.15) *is satisfied for the* (*RQ*) *problem* (2.1). *Consider the following nonconvex homogeneous minimization problem*:

$$\varphi^* \equiv \min_{\mathbf{z} \in \mathbb{R}^n, s \in \mathbb{R}} \{\varphi_1(\mathbf{z}, s) : \varphi_2(\mathbf{z}, s) = 1, \varphi_3(\mathbf{z}, s) \leq 0\}, \qquad (3.1)$$

where

$$\varphi_i(\mathbf{z}, s) = \mathbf{z}^T \mathbf{A}_i \mathbf{z} + 2\mathbf{b}_i^T \mathbf{z}s + c_i s^2, \quad i = 1, 2,$$
$$\varphi_3(\mathbf{z}, s) = \|\mathbf{L}\mathbf{z}\|^2 - \rho s^2.$$

Then $\varphi^* = f^*$. *Moreover, any optimal solution* $(\mathbf{z}^*, s^*)$ *of* (3.1) *satisfies* $s^* \neq 0$ *and* $\mathbf{x}^* = \frac{1}{s^*}\mathbf{z}^*$ *is an optimal solution of the RQ problem.*

Proof By using the change of variables $\mathbf{x} = \mathbf{y}/t, t \neq 0$, the RQ problem transforms into the following equivalent minimization problem:

$$f^* = \min_{\mathbf{y} \in \mathbb{R}^n, t \neq 0} \left\{ \frac{\varphi_1(\mathbf{y}, t)}{\varphi_2(\mathbf{y}, t)} : \varphi_3(\mathbf{y}, t) \leq 0 \right\}. \qquad (3.2)$$

Recall that by Assumption 1 it follows that $\varphi_2(\mathbf{y}, t) > 0$ for every $\mathbf{y} \in \mathbb{R}^n, t \neq 0$ such that $\varphi_3(\mathbf{y}, t) \leq 0$ so that problem (3.2) is well-defined. Now, consider the following problem:

$$\varphi^* = \min_{\mathbf{z} \in \mathbb{R}^n, s \neq 0} \{\varphi_1(\mathbf{z}, s) : \varphi_2(\mathbf{z}, s) = 1, \varphi_3(\mathbf{z}, s) \leq 0\}. \qquad (3.3)$$

First, we show that the feasible set of problem (3.3) is compact, so that the minimum is attained. For that, note that for any $\eta \geq 0$, the following inclusion obviously holds for the feasible set of (3.3):

$$\{(\mathbf{z}, s) : \varphi_2(\mathbf{z}, s) = 1, \varphi_3(\mathbf{z}, s) \leq 0\} \subseteq \{(\mathbf{z}, s) : \varphi_2(\mathbf{z}, s) + \eta \varphi_3(\mathbf{z}, s) \leq 1\}.$$

Therefore, under Assumption 1, it follows that the set in the right-hand-side of the latter inclusion describes a non-degenerate ellipsoid, and thus is compact, and hence so is the feasible set of problem (3.3).

We will now show that $\varphi^* = f^*$. Indeed, suppose that $(\mathbf{y}^*, t^*)$ with $t^* \neq 0$ is an optimal solution of (3.2), then $(\mathbf{z}^*, s^*)$ defined by

$$\mathbf{z}^* = \frac{1}{\sqrt{\varphi_2(\mathbf{y}^*, t^*)}} \mathbf{y}^*, \quad s^* = \frac{1}{\sqrt{\varphi_2(\mathbf{y}^*, t^*)}} t^*$$

is a feasible point of (3.3) since, by using the fact that φ_i are homogeneous functions, we have

$$\varphi_3(\mathbf{z}^*, s^*) = \frac{1}{\varphi_2(\mathbf{y}^*, t^*)} \varphi_3(\mathbf{y}^*, t^*) \leq 0,$$

$$\varphi_2(\mathbf{z}^*, s^*) = \frac{1}{\varphi_2(\mathbf{y}^*, t^*)} \varphi_2(\mathbf{y}^*, t^*) = 1$$

and thus

$$\varphi^* \leq \varphi_1(\mathbf{z}^*, s^*) = \frac{\varphi_1(\mathbf{y}^*, t^*)}{\varphi_2(\mathbf{y}^*, t^*)} = f^*.$$

On the other hand, let $(\mathbf{z}^*, s^*)$ be an optimal solution of (3.3). Then $(\mathbf{z}^*, s^*)$ is obviously also a feasible solution of (3.2) and thus

$$f^* \leq \frac{\varphi_1(\mathbf{z}^*, s^*)}{\varphi_2(\mathbf{z}^*, s^*)} = \varphi_1(\mathbf{z}^*, s^*) = \varphi^*.$$

We would like to show now that the constraint $s \neq 0$ in problem (3.3) can be omitted. Suppose on the contrary that for some optimal solution $(\mathbf{z}^*, s^*)$ one has $s^* = 0$. Substituting $s = 0$ in the minimization problem (3.3) we have:

$$f^* = \varphi^* = \min_{\mathbf{z} \in \mathbb{R}^n} \{\varphi_1(\mathbf{z}, 0) : \varphi_2(\mathbf{z}, 0) = 1, \varphi_3(\mathbf{z}, 0) \leq 0\}$$

$$= \min_{\mathbf{z} \in \mathbb{R}^n} \left\{ \mathbf{z}^T \mathbf{A}_1 \mathbf{z} : \mathbf{z}^T \mathbf{A}_2 \mathbf{z} = 1, \mathbf{L}\mathbf{z} = \mathbf{0} \right\}. \tag{3.4}$$

In the case where $r = n$, the minimization problem is infeasible—a contradiction to the fact that $f^* \neq \infty$; in the case where $r < n$, we can replace the constraint $\mathbf{L}\mathbf{z} = \mathbf{0}$ with the linear relation $\mathbf{z} = \mathbf{F}\mathbf{u}$ where $\mathbf{u} \in \mathbb{R}^{n-r}$ and (3.4) is transformed to

$$f^* = \min_{\mathbf{u} \in \mathbb{R}^{n-r}} \{\mathbf{u}^T \mathbf{F}^T \mathbf{A}_1 \mathbf{F}\mathbf{u} : \mathbf{u}^T \mathbf{F}^T \mathbf{A}_2 \mathbf{F}\mathbf{u} = 1\} = \lambda_{\min}(\mathbf{F}^T \mathbf{A}_1 \mathbf{F}, \mathbf{F}^T \mathbf{A}_2 \mathbf{F}),$$

which contradicts (2.11). □

3.2 Strong duality of the non-convex quadratic formulation

Thanks to Lemma 3 we have just shown that solving the RQ problem amounts to solving a nonconvex homogeneous quadratic problem (3.1). However, we will show that problem (3.1) admits an exact semidefinite reformulation by applying a strong duality result for homogeneous nonconvex quadratic problems with two quadratic constraints that was first proven in [17]. For more recent and related results see [23], and references therein.

Proposition 2 ([17, Proposition 4.1]) *Consider the following homogeneous nonconvex quadratic problem*:

$$(H) \quad \inf\{\mathbf{y}^T \mathbf{R}_1 \mathbf{y} : \mathbf{y}^T \mathbf{R}_2 \mathbf{y} = a_2, \mathbf{y}^T \mathbf{R}_3 \mathbf{y} \leq a_3\},$$

where $\mathbf{R}_i$ are $n \times n$ symmetric matrices ($i = 1, 2, 3$) with $n \geq 3$, a_2 and a_3 be real numbers such that $a_2 \neq 0$. Suppose that the following two conditions are satisfied:

A. *there exist $\mu_2, \mu_3 \in \mathbb{R}$ such that $\mu_2 \mathbf{R}_2 + \mu_3 \mathbf{R}_3 \succ \mathbf{0}$,*
B. *there exists $\tilde{\mathbf{y}} \in \mathbb{R}^n$ such that $\tilde{\mathbf{y}}^T \mathbf{R}_2 \tilde{\mathbf{y}} = a_2$, $\tilde{\mathbf{y}}^T \mathbf{R}_3 \tilde{\mathbf{y}} < a_3$.*

Then $val(H) = val(DH)$ where (DH) is the dual problem

$$(DH) \quad \sup_{\beta \geq 0, \alpha} \{\alpha a_2 - \beta a_3 : \mathbf{R}_1 \succeq \alpha \mathbf{R}_2 - \beta \mathbf{R}_3\}.$$

Combining the above result, together with Lemma 3, we are now able to establish the promised semidefinite reformulation of problem (3.1).

Theorem 2 *Let $n \geq 2$ and suppose that condition (2.15) is satisfied. Then $val(D) = f^*$ where (D) is given by*

$$(D) : \quad \max_{\beta \geq 0, \alpha} \left\{ \alpha : \begin{pmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^T & c_1 \end{pmatrix} \succeq \alpha \begin{pmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^T & c_2 \end{pmatrix} - \beta \begin{pmatrix} \mathbf{L}^T \mathbf{L} & \mathbf{0} \\ \mathbf{0} & -\rho \end{pmatrix} \right\}.$$

Proof Let

$$\mathbf{R}_i = \begin{pmatrix} \mathbf{A}_i & \mathbf{b}_i \\ \mathbf{b}_i^T & c_i \end{pmatrix}, \mathbf{R}_3 = \begin{pmatrix} \mathbf{L}^T \mathbf{L} & \mathbf{0} \\ \mathbf{0}^T & -\rho \end{pmatrix}, a_2 = 1, a_3 = 0 \quad i = 1, 2.$$

Then the pair of problems (H) and (DH) reduce to (3.1) and its dual (D) respectively. Conditions A and B of Proposition 2 are satisfied for $\mu_2 = 1, \mu_3 = \eta$ (see LMI (2.2)) and with a vector $\tilde{\mathbf{y}} = (\mathbf{0}; 1/\sqrt{c_2})$, where $c_2 > 0$ by Proposition 1(i). Therefore, $val(3.1) = val(D)$, and by invoking Lemma 3 it follows that $f^* = val(D)$. □

From Theorem 2 it follows that we can find the optimal *function value* of the nonconvex problem (RQ) by solving the convex SDP problem (D). Furthermore, it is interesting to note that we can also recover a global optimal solution $\mathbf{x}^*$ of problem

(RQ) from an optimal solution (α^*, β^*) of the dual problem (D). Indeed, we remark that x^* is an optimal solution of problem (RQ) if and only if

$$\mathbf{x}^* \in \underset{\mathbf{x} \in \mathbb{R}^n}{\operatorname{argmin}} \{ f_2(\mathbf{x})(f(\mathbf{x}) - f^*) : \|\mathbf{L}\mathbf{x}\|^2 \le \rho \}, \tag{3.5}$$

which is just a generalized trust region subproblem for which several efficient algorithms exist (see the discussion below at the beginning of Sect. 4). Thus, since $f^* = \alpha^*$, a global optimal solution for (RQ) can be derived via the following procedure:

Procedure RQ-SDP

1. Solve the SDP problem (D) and obtain a solution (α^*, β^*).
2. Set $\mathbf{x}^*$ to be an optimal solution of the problem (3.5) with $f^* = \alpha^*$, i.e.,

$$\min_{\mathbf{x} \in \mathbb{R}^n} \{ \mathbf{x}^T (\mathbf{A}_1 - \alpha^* \mathbf{A}_2)\mathbf{x} + 2(\mathbf{b}_1 - \alpha^* \mathbf{b}_2)^T \mathbf{x} + c_1 - \alpha^* c_2 : \|\mathbf{L}\mathbf{x}\|^2 \le \rho \}.$$

4 Convergence analysis of a fixed point algorithm

Building on the hidden convexity properties we have derived in the previous sections for problem (RQ), in this section we analyze the iterative scheme recently proposed by Sima et al. [20]. The iterative scheme involves the solution of a GTRS problem of the form

$$\min \{ \mathbf{x}^T \mathbf{A} \mathbf{x} + 2\mathbf{b}^T \mathbf{x} : \|\mathbf{L}\mathbf{x}\|^2 \le \rho \}$$

at each iteration. There are several methods for solving GTRS problems. One approach is to transform the problem into one of solving a single variable secular equation [5]. Another approach, used in [20], is to formulate the problem as a quadratic eigenvalue problem, for which efficient algorithms are known to exist. For medium and large-scale problems one can use a modification of Moré and Sorensen's method [15] or algorithms based on Krylov subspace methods such as those devised in [21]. For a related work on parametric eigenvalue problems see [19].

It was observed in [20] that the iterative scheme is robust with respect to the initial vector, which indicates—although not proven—that the method converges to a *global* minimum. The latter claim was validated numerically in [5]. Moreover, it has been observed that the method in [20] converges at a very fast rate; specifically, it requires to solve no more than 5 GTRS problems, independently of the dimension of the problem. In this section, a *global* convergence analysis of this iterative scheme is developed for the more general problem (RQ) and by so doing, this also provides a theoretical justification that supports the excellent performance of the method given in [20].

4.1 Fixed point iterations

The starting point is the observation already mentioned in (3.5) that $\mathbf{x}^*$ is an optimal solution of problem (RQ) given in (2.1) if and only if

$$\mathbf{x}^* \in \underset{\mathbf{y} \in \mathbb{R}^n}{\operatorname{argmin}} \{f_2(\mathbf{y})(f(\mathbf{y}) - f(\mathbf{x}^*)) \ : \ \|\mathbf{L}\mathbf{y}\|^2 \leq \rho\}, \tag{4.1}$$

which naturally leads to consider the following fixed point iterations:

$$\mathbf{x}_{k+1} \in \underset{\mathbf{y} \in \mathbb{R}^n}{\operatorname{argmin}} \{f_2(\mathbf{y})(f(\mathbf{y}) - f(\mathbf{x}_k)) : \|\mathbf{L}\mathbf{y}\|^2 \leq \rho\}. \tag{4.2}$$

Choice of initial vector. In the case $r = n$, $\mathbf{x}_0$ can be chosen as an arbitrary feasible vector (for example, $\mathbf{x}_0 = 0$). In the case $r < n$ the vector $\mathbf{x}_0$ will be chosen to be equal to $\mathbf{F}\mathbf{v}/t$, where $(\mathbf{v}; t)$ is a minimum generalized eigenvector of $\mathbf{M}_1, \mathbf{M}_2$, and the matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ are defined in (2.10). Recall that by the first part of Theorem 1, $t \neq 0$. Moreover, $\mathbf{x}_0 = \mathbf{F}\mathbf{v}/t$ is a feasible point of problem (RQ) and under our blanket hypothesis,

$$f(\mathbf{x}_0) = \lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) < \lambda_{\min}(\mathbf{F}^T \mathbf{A}_1 \mathbf{F}, \mathbf{F}^T \mathbf{A}_2 \mathbf{F}).$$

To guarantee that the sequence $\{\mathbf{x}_k\}$ given by (4.2) is well-defined, we need to show that the minimum in problem (4.2) is finite and attained. Let γ be some upper bound on the value of problem (4.2). For any $k \geq 0$ consider the set S_k defined by

$$S_k := \{\mathbf{y} \in \mathbb{R}^n \mid f_1(\mathbf{y}) - f(\mathbf{x}_k) f_2(\mathbf{y}) \leq \gamma, \|\mathbf{L}\mathbf{y}\|^2 \leq \rho\}.$$

Since S_k is an intersection of a nonempty level set of the objective function in (4.2) (recall that $f_2(\mathbf{y})(f(\mathbf{y}) - f(\mathbf{x}_k)) = f_1(\mathbf{y}) - f(\mathbf{x}_k) f_2(\mathbf{y})$) and the feasible set of (4.2), it follows that problem (4.2) is the same as $\min\{f_1(\mathbf{y}) - f(\mathbf{x}_k) f_2(\mathbf{y}) : \mathbf{y} \in S_k\}$. Therefore, to show that the minimum in (4.2) is finite and attained it is sufficient to show that the set S_k is bounded. Invoking Lemma 2(ii), it follows that S_k is bounded if the following condition holds:

$$f(\mathbf{x}_k) < \lambda_{\min}(\mathbf{F}^T \mathbf{A}_1 \mathbf{F}, \mathbf{F}^T \mathbf{A}_2 \mathbf{F}). \tag{4.3}$$

Since $\mathbf{x}_0$ was chosen to satisfy (4.3), all that is left to prove in order to show that the iterations (4.2) are well-defined is that the sequence is nonincreasing.

Lemma 4 *Let $\{\mathbf{x}_k\}$ be the sequence generated by (4.2). Then $f(\mathbf{x}_{k+1}) \leq f(\mathbf{x}_k)$ for every $k \geq 0$.*

Proof Using the fact that $\mathbf{x}_{k+1}$ is a minimizer of problem (4.2) we have

$$f_2(\mathbf{x}_{k+1})(f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)) \leq f_2(\mathbf{x}_k)(f(\mathbf{x}_k) - f(\mathbf{x}_k)) = 0.$$

The monotonicity property follows from the fact that $f_2(\mathbf{x}_{k+1}) > 0$. □

4.2 Convergence analysis

We now analyze the basic iteration scheme (4.2). The next result shows that the sequence of function values converges to the *global* optimal solution with a linear rate.

Theorem 3 *Let* $\{\mathbf{x}_k\}$ *be the sequence generated by* (4.2) *and* $\mathbf{x}^*$ *be an optimal solution of the RQ problem* (2.1). *Then*:

(i) *There exists* U *such that*

$$\|\mathbf{x}_k\| \leq U, \quad k = 0, 1, \ldots \tag{4.4}$$

(ii) *The sequence of function values converge to the optimal value* $f(\mathbf{x}^*)$ *with a linear rate of convergence: there exists* $\gamma \in (0, 1)$ *such that*

$$f(\mathbf{x}_{k+1}) - f(\mathbf{x}^*) \leq \gamma(f(\mathbf{x}_k) - f(\mathbf{x}^*)) \quad \text{for every } k \geq 0, \tag{4.5}$$

where

$$\gamma = 1 - \frac{\delta}{\lambda_{\max}(\mathbf{A}_2)U^2 + 2\|\mathbf{b}_2\|U + c_2}, \tag{4.6}$$

and δ *is given in* (2.4).

Proof (i) If $r = n$ then the result follows from the compactness of the feasible set. Otherwise, if $r < n$, we note that the sequence $\{f(\mathbf{x}_k)\}$ is monotonically decreasing and we therefore conclude that the following inequalities hold true for every k:

$$f(\mathbf{x}_k) \leq \alpha,$$
$$\|\mathbf{L}\mathbf{x}_k\|^2 \leq \rho.$$

where $\alpha = \lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2)$. Since here $\alpha = \lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) < \lambda_{\min}(\mathbf{F}^T\mathbf{A}_1\mathbf{F}, \mathbf{F}^T\mathbf{A}_2\mathbf{F})$, by Lemma 2(ii), it follows that the sequence $\{\mathbf{x}_k\}$ is bounded, which proves (i).

(ii) By the definition of $\mathbf{x}_{k+1}$ we have

$$f_2(\mathbf{x}_{k+1})(f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k)) \leq f_2(\mathbf{x}^*)(f(\mathbf{x}^*) - f(\mathbf{x}_k))$$

and thus

$$f(\mathbf{x}_{k+1}) - f(\mathbf{x}^*) \leq \frac{f_2(\mathbf{x}^*)}{f_2(\mathbf{x}_{k+1})}(f(\mathbf{x}^*) - f(\mathbf{x}_k)) - (f(\mathbf{x}^*) - f(\mathbf{x}_k))$$

$$\leq \left(1 - \frac{f_2(\mathbf{x}^*)}{f_2(\mathbf{x}_{k+1})}\right)(f(\mathbf{x}_k) - f(\mathbf{x}^*)). \tag{4.7}$$

Using (4.4) it is easy to see that

$$f_2(\mathbf{x}_{k+1}) \leq \lambda_{\max}(\mathbf{A}_2)U^2 + 2\|\mathbf{b}_2\|U + c_2 \tag{4.8}$$

and by Proposition 1 (iii) we have that $f_2(\mathbf{x}^*) \geq \delta > 0$ with δ is given in (2.4). Therefore, from (4.7) the desired result (4.5) follows with γ given in (4.6). $\square$

Remark 2 The upper bound U can be explicitly computed in terms of the problem's data, see Lemma A in the Appendix.

Inequality (4.5) implies that the following inequality holds true:

$$f(\mathbf{x}_k) - f(\mathbf{x}^*) \leq \gamma^k(f(\mathbf{x}_0) - f(\mathbf{x}^*)). \tag{4.9}$$

The next result will provide a convergence rate result for the sequence $\{\mathbf{x}_k\}$. For that purpose we first recall the well known optimality conditions for the GTRS problem, (see [15]): $\mathbf{x}^*$ is an optimal solution of the problem $\min\{q(\mathbf{x}) : \|\mathbf{L}\mathbf{x}\|^2 \leq \rho\}$, where $q : \mathbb{R}^n \to \mathbb{R}$ is a quadratic function if and only if there exists $\lambda^* \geq 0$ such that

$$\begin{aligned}
\nabla q(\mathbf{x}^*) + \lambda^* \mathbf{L}^T \mathbf{L}\mathbf{x}^* &= \mathbf{0}, \\
\nabla^2 q(\mathbf{x}^*) + \lambda^* \mathbf{L}^T \mathbf{L} &\succeq \mathbf{0}, \\
\|\mathbf{L}\mathbf{x}^*\|^2 &\leq \rho, \\
\lambda^*(\|\mathbf{L}\mathbf{x}^*\|^2 - \rho) &= 0.
\end{aligned} \tag{4.10}$$

Furthermore, if (4.10) is replaced with

$$\nabla^2 q(\mathbf{x}^*) + \lambda^* \mathbf{L}^T \mathbf{L} \succ \mathbf{0}, \tag{4.11}$$

then the optimal solution is unique.

Theorem 4 *Let $\{\mathbf{x}_k\}$ be the sequence generated by (4.2) and $\mathbf{x}^*$ be an optimal solution of the RQ problem (2.1). Suppose that condition (4.11) is satisfied for problem (4.1). Then*

$$\|\mathbf{x}_k - \mathbf{x}^*\|^2 \leq C(f(\mathbf{x}_k) - f(\mathbf{x}^*)) \quad \text{for every } k \geq 0, \tag{4.12}$$

where

$$C = 2\frac{\lambda_{\max}(\mathbf{A}_2)U^2 + 2\|\mathbf{b}_2\|U + c_2}{\lambda_{\min}(2\mathbf{A}_1 - 2f(\mathbf{x}^*)\mathbf{A}_2 + \lambda^*\mathbf{L}^T\mathbf{L})} \tag{4.13}$$

In particular, the sequence $\{\mathbf{x}_k\}$ converges at a linear rate to a unique optimal solution $\mathbf{x}^*$ of problem (RQ).

Proof Define

$$g(\mathbf{x}) = f_1(\mathbf{x}) - f(\mathbf{x}^*)f_2(\mathbf{x}). \tag{4.14}$$

Then problem (4.1) is the same as

$$\min_{\mathbf{x} \in \mathbb{R}^n} \{g(\mathbf{x}) : \|\mathbf{L}\mathbf{x}\|^2 \leq \rho\}. \tag{4.15}$$

Let $\lambda^* \geq 0$ be the corresponding optimal Lagrange multiplier for problem (4.15). Since

$$g(\mathbf{x}_k) - g(\mathbf{x}^*) = (\mathbf{x}_k - \mathbf{x}^*)^T \nabla g(\mathbf{x}^*) + \frac{1}{2}(\mathbf{x}_k - \mathbf{x}^*)^T \nabla^2 g(\mathbf{x}^*)(\mathbf{x}_k - \mathbf{x}^*),$$

then using (4.10) we obtain

$$
\begin{aligned}
g(\mathbf{x}_k) - g(\mathbf{x}^*) &= (\mathbf{x}_k - \mathbf{x}^*)^T \underbrace{(\nabla g(\mathbf{x}^*) + \lambda^* \mathbf{L}^T \mathbf{L} \mathbf{x}^*)}_{0} \\
&\quad + \frac{1}{2}(\mathbf{x}_k - \mathbf{x}^*)^T (\nabla^2 g(\mathbf{x}^*) + \lambda^* \mathbf{L}^T \mathbf{L})(\mathbf{x}_k - \mathbf{x}^*) \\
&\quad - \lambda^* \left((\mathbf{x}_k - \mathbf{x}^*)^T \mathbf{L}^T \mathbf{L} \mathbf{x}^* + \frac{1}{2}(\mathbf{x}_k - \mathbf{x}^*)\mathbf{L}^T \mathbf{L}(\mathbf{x}_k - \mathbf{x}^*) \right) \\
&= \frac{1}{2}(\mathbf{x}_k - \mathbf{x}^*)^T (\nabla^2 g(\mathbf{x}^*) + \lambda^* \mathbf{L}^T \mathbf{L})(\mathbf{x}_k - \mathbf{x}^*) + \frac{\lambda^*}{2}\left(\|\mathbf{L}\mathbf{x}^*\|^2 - \|\mathbf{L}\mathbf{x}_k\|^2 \right) \\
&\geq \frac{1}{2}(\mathbf{x}_k - \mathbf{x}^*)^T (\nabla^2 g(\mathbf{x}^*) + \lambda^* \mathbf{L}^T \mathbf{L})(\mathbf{x}_k - \mathbf{x}^*) + \frac{\lambda^*}{2}\left(\|\mathbf{L}\mathbf{x}^*\|^2 - \rho \right) \\
&= \frac{1}{2}(\mathbf{x}_k - \mathbf{x}^*)^T (\nabla^2 g(\mathbf{x}^*) + \lambda^* \mathbf{L}^T \mathbf{L})(\mathbf{x}_k - \mathbf{x}^*) \\
&\geq \frac{1}{2}\lambda_{\min}(\nabla^2 g(\mathbf{x}^*) + \lambda^* \mathbf{L}^T \mathbf{L})\|\mathbf{x}_k - \mathbf{x}^*\|^2 \\
&= \frac{1}{2}\lambda_{\min}(2\mathbf{A}_1 - 2f(\mathbf{x}^*)\mathbf{A}_2 + \lambda^* \mathbf{L}^T \mathbf{L})\|\mathbf{x}_k - \mathbf{x}^*\|^2. \qquad (4.16)
\end{aligned}
$$

Using (4.14) one has the relation

$$
f(\mathbf{x}_k) - f(\mathbf{x}^*) = \frac{1}{f_2(\mathbf{x}_k)}(g(\mathbf{x}_k) - g(\mathbf{x}^*)),
$$

which, together with (4.8) and (4.16), implies that (4.12) is satisfied with C given in (4.13). Finally,

$$
\|\mathbf{x}_k - \mathbf{x}^*\| \overset{(4.12)}{\leq} \sqrt{C(f(\mathbf{x}_k) - f(\mathbf{x}^*))} \overset{(4.9)}{\leq} \sqrt{C(f(\mathbf{x}_0) - f(\mathbf{x}^*))}\gamma^{k/2}, \qquad (4.17)
$$

where γ is given in (4.6), which proves the last claim.   □

Our last result shows that

(i)   the linear convergence rate result for function values established in Theorem 3 can be improved to superlinear, and

(ii) the dependency of the number of iterations on the optimality tolerance ε grows as $O(\sqrt{\ln(1/\varepsilon)})$.

Theorem 5 *Let $\{\mathbf{x}_k\}$ be the sequence generated by (4.2) and $\mathbf{x}^*$ be an optimal solution of the RQ problem (2.1). Suppose that condition (4.11) is satisfied for problem (4.1). Then the corresponding sequence of function values $\{f(\mathbf{x}_k)\}$ converges at least superlinearly. Moreover, an ε-global optimal solution, i.e., satisfying $f(\mathbf{x}_n) - f(\mathbf{x}^*) \leq \varepsilon$, is reached after at most $A + B\sqrt{\ln(1/\varepsilon)}$ iterations, for some positive constants A, B.*

Proof Note that by the mean value theorem we have

$$
\begin{aligned}
|f_2(\mathbf{x}_k) - f_2(\mathbf{x}^*)| &= |\nabla f_2(\omega \mathbf{x}_k + (1-\omega)\mathbf{x}^*)^T (\mathbf{x}_k - \mathbf{x}^*)| \\
&\leq 2\|\mathbf{A}_2(\omega \mathbf{x}_k + (1-\omega)\mathbf{x}^*) + \mathbf{b}_2\|\|\mathbf{x}_k - \mathbf{x}^*\| \\
&\leq 2(\|\mathbf{A}_2\|U + \|\mathbf{b}_2\|)\|\mathbf{x}_k - \mathbf{x}^*\|,
\end{aligned}
$$

where the last equation follows from Theorem 3(i).
Now,

$$
\frac{f(\mathbf{x}_{k+1}) - f(\mathbf{x}^*)}{f(\mathbf{x}_k) - f(\mathbf{x}^*)} \overset{(4.7)}{\leq} \frac{f_2(\mathbf{x}_{k+1}) - f_2(\mathbf{x}^*)}{f_2(\mathbf{x}_{k+1})} \leq \frac{2(\|\mathbf{A}_2\|U + \|\mathbf{b}_2\|)}{\delta}\|\mathbf{x}_{k+1} - \mathbf{x}^*\|
$$

$$
\overset{(4.16)}{\leq} \frac{2(\|\mathbf{A}_2\|U + \|\mathbf{b}_2\|)}{\delta}\sqrt{C(f(\mathbf{x}_0) - f(\mathbf{x}^*))}\gamma^{k/2},
$$

where $\gamma \in (0, 1)$ is given in (4.6). Thus, with

$$
D = \frac{2(\|\mathbf{A}_2\|U + \|\mathbf{b}_2\|)}{\delta}\sqrt{C(f(\mathbf{x}_0) - f(\mathbf{x}^*))},
$$

we have proven that

$$
\frac{f(\mathbf{x}_{k+1}) - f(\mathbf{x}^*)}{f(\mathbf{x}_k) - f(\mathbf{x}^*)} \leq D\gamma^{k/2}. \tag{4.18}
$$

Since $D\gamma^{k/2} \to 0$, the superlinear convergence rate of function values follows. To derive the second claim of the theorem, we proceed as follows. Using (4.18), we have,

$$
\frac{f(\mathbf{x}_n) - f(\mathbf{x}^*)}{f(\mathbf{x}_0) - f(\mathbf{x}^*)} = \prod_{k=0}^{n-1} \frac{f(\mathbf{x}_{k+1}) - f(\mathbf{x}^*)}{f(\mathbf{x}_k) - f(\mathbf{x}^*)} \leq \prod_{k=0}^{n-1} D\gamma^{k/2} = D^n \gamma^{(n^2-n)/4},
$$

and hence,

$$
f(\mathbf{x}_n) - f(\mathbf{x}^*) \leq (f(\mathbf{x}_0) - f(\mathbf{x}^*))D^n \gamma^{(n^2-n)/4} = \gamma^{n^2/4 + \alpha_1 n + \alpha_2},
$$

where $\alpha_1 = -1/4 + \ln_\gamma D$, $\alpha_2 = \ln_\gamma(f(\mathbf{x}_0) - f(\mathbf{x}^*))$. Now, the inequality

$$
\gamma^{n^2/4 + \alpha_1 n + \alpha_2} \leq \varepsilon
$$

holds true if and only if

$$
n^2/4 + \alpha_1 n + \alpha_2 \geq -\ln_\gamma(1/\varepsilon).
$$

Let $\Delta = \alpha_1^2 - \alpha_2 - \ln_\gamma(1/\varepsilon)$. Then the last inequality is obviously satisfied for any n whenever $\Delta < 0$ and for $n > 2(-\alpha_1 + \sqrt{\Delta})$ if $\Delta \geq 0$. It readily follows that if

$n > 2|\alpha_1| + 2\sqrt{|\Delta|}$ then $f(\mathbf{x}_n) - f(\mathbf{x}^*) \leq \varepsilon$. Therefore, since

$$|\alpha_1| + \sqrt{|\Delta|} \leq |\alpha_1| + \sqrt{|\alpha_1^2| + |\alpha_2| - \ln_\gamma(1/\varepsilon)} \leq 2|\alpha_1| + \sqrt{|\alpha_2|} + \sqrt{-\frac{1}{\ln\gamma}}\sqrt{\ln(1/\varepsilon)},$$

the desired result follows, with

$$A = |4\ln_\gamma D - 1| + 2\sqrt{|\ln_\gamma(f(\mathbf{x}_0) - f(\mathbf{x}^*))|}; \quad B = \sqrt{-\frac{4}{\ln(\gamma)}}.$$

$\square$

Finally, as a consequence of Theorem 5 and (4.12) we obtain,

Corollary 1 *Under the assumptions of Theorem 4.3, the number of iterations required to reach a point $\mathbf{x}_k$ such that $\|\mathbf{x}_k - \mathbf{x}^*\| \leq \varepsilon$ grows as $O(\sqrt{\ln(1/\varepsilon)})$.*

A Appendix

We show that the upper bound U given in Theorem 3(i) can be computed in terms of the problem's data.

Lemma A *Let $\{\mathbf{x}_k\}$ be the sequence generated by (4.2) and let $\mathbf{F} \in \mathbb{R}^{n \times (n-r)}$ be a matrix whose columns form an orthonormal basis for the null space of $\mathbf{L}$. Then $\|\mathbf{x}_k\| \leq U$ for all $k = 0, \ldots,$ where U is determined as follows*:

$$U = \begin{cases} \sqrt{\dfrac{\rho}{\lambda_{\min}(\mathbf{L}^T\mathbf{L})}} & \text{if } r = n \\ 2\sqrt{\dfrac{U_1^2}{\lambda_{\min}(\tilde{\mathbf{A}})^2} + \dfrac{U_2}{\lambda_{\min}(\tilde{\mathbf{A}})}} + \|\mathbf{L}\|\dfrac{\sqrt{\rho}}{\lambda_{\min}(\mathbf{L}\mathbf{L}^T)} & \text{if } r < n \end{cases}$$

where

$$U_1 = \frac{\|\mathbf{A}_1 - \alpha\mathbf{A}_2\|\|\mathbf{L}\|\sqrt{\rho}}{\lambda_{\min}(\mathbf{L}\mathbf{L}^T)} + \|\mathbf{b}_1 - \alpha\mathbf{b}_2\|, \tag{A.1}$$

$$U_2 = \frac{\|\mathbf{L}(\mathbf{A}_1 - \alpha\mathbf{A}_2)\mathbf{L}^T\|\rho}{\lambda_{\min}(\mathbf{L}\mathbf{L}^T)^2} + 2\frac{\|\mathbf{b}_1 - \alpha\mathbf{b}_2\|\|\mathbf{L}\|\sqrt{\rho}}{\lambda_{\min}(\mathbf{L}\mathbf{L}^T)} + |c_1 - \alpha c_2|, \tag{A.2}$$

$$\tilde{\mathbf{A}} = \mathbf{F}^T(\mathbf{A}_1 - \alpha\mathbf{A}_2)\mathbf{F}.$$

Proof If $r = n$ then the feasibility constraint $\|\mathbf{L}\mathbf{x}_k\|^2 \leq \rho$ implies that $\|\mathbf{x}_k\|^2 \leq \rho/\lambda_{\min}(\mathbf{L}^T\mathbf{L})$. Otherwise, when $r < n$, we can consider the following decomposition:

$$\mathbf{x}_k = \mathbf{F}\mathbf{v}_k + \mathbf{L}^T\mathbf{w}_k, \tag{A.3}$$

Plugging (A.3) into the feasibility constraint we obtain:

$$\|\mathbf{L}\mathbf{L}^T\mathbf{w}_k\|^2 \leq \rho,$$

which yields:

$$\|\mathbf{w}_k\|^2 \leq \rho/\lambda_{\min}(\mathbf{L}\mathbf{L}^T)^2. \tag{A.4}$$

We continue with bounding $\mathbf{v}_k$. In order to so, we recall that the sequence $\mathbf{x}_k$ satisfies

$$f(\mathbf{x}_k) \leq \alpha, \tag{A.5}$$

where $\alpha = \lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) < \lambda_{\min}(\mathbf{F}^T\mathbf{A}_1\mathbf{F}, \mathbf{F}^T\mathbf{A}_2\mathbf{F})$. This also implies that $\mathbf{F}^T(\mathbf{A}_1 - \alpha\mathbf{A}_2)\mathbf{F} \succ \mathbf{0}$. Now, plugging the decomposition (A.3) into (A.5) results in

$$(\mathbf{F}\mathbf{v}_k + \mathbf{L}^T\mathbf{w}_k)^T(\mathbf{A}_1 - \alpha\mathbf{A}_2)(\mathbf{F}\mathbf{v}_k + \mathbf{L}^T\mathbf{w}_k)$$
$$+2(\mathbf{b}_1 - \alpha\mathbf{b}_2)^T(\mathbf{F}\mathbf{v}_k + \mathbf{L}^T\mathbf{w}_k) + c_1 - \alpha c_2 \leq 0,$$

which can also be written as

$$\mathbf{v}_k^T\widetilde{\mathbf{A}}\mathbf{v}_k + 2\tilde{\mathbf{b}}^T\mathbf{v}_k + \tilde{c} \leq 0, \tag{A.6}$$

where

$$\widetilde{\mathbf{A}} = \mathbf{F}^T(\mathbf{A}_1 - \alpha\mathbf{A}_2)\mathbf{F},$$
$$\tilde{\mathbf{b}} = \mathbf{F}^T(\mathbf{A}_1 - \alpha\mathbf{A}_2)\mathbf{L}^T\mathbf{w}_k + \mathbf{F}^T(\mathbf{b}_1 - \alpha\mathbf{b}_2),$$
$$\tilde{c} = \mathbf{w}_k^T\mathbf{L}(\mathbf{A}_1 - \alpha\mathbf{A}_2)\mathbf{L}^T\mathbf{w}_k + 2(\mathbf{b}_1 - \alpha\mathbf{b}_2)^T\mathbf{L}^T\mathbf{w}_k + c_1 - \alpha c_2.$$

A direct computation shows that (A.6) is equivalent to

$$\|\mathbf{v}_k + \widetilde{\mathbf{A}}^{-1}\mathbf{b}\|^2 \leq \frac{\tilde{\mathbf{b}}^T\widetilde{\mathbf{A}}^{-1}\tilde{\mathbf{b}} - \tilde{c}}{\lambda_{\min}(\widetilde{\mathbf{A}})},$$

which implies:

$$\|\mathbf{v}_k\| \leq \sqrt{\frac{\tilde{\mathbf{b}}^T\widetilde{\mathbf{A}}^{-1}\tilde{\mathbf{b}} - \tilde{c}}{\lambda_{\min}(\widetilde{\mathbf{A}})}} + \|\widetilde{\mathbf{A}}^{-1}\tilde{\mathbf{b}}\|$$

$$\leq \sqrt{\frac{\|\tilde{\mathbf{b}}\|^2}{\lambda_{\min}(\widetilde{\mathbf{A}})^2} + \frac{|\tilde{c}|}{\lambda_{\min}(\widetilde{\mathbf{A}})}} + \frac{\|\tilde{\mathbf{b}}\|}{\lambda_{\min}(\widetilde{\mathbf{A}})}$$

$$\leq 2\sqrt{\frac{\|\tilde{\mathbf{b}}\|^2}{\lambda_{\min}(\widetilde{\mathbf{A}})^2} + \frac{|\tilde{c}|}{\lambda_{\min}(\widetilde{\mathbf{A}})}}.$$

We now compute upper bounds on $\|\tilde{\mathbf{b}}\|$ and $|\tilde{c}|$:

$$\|\tilde{\mathbf{b}}\| \leq \|\mathbf{A}_1 - \alpha\mathbf{A}_2\|\|\mathbf{L}\|\|\mathbf{w}_k\| + \|\mathbf{b}_1 - \alpha\mathbf{b}_2\|, \tag{A.7}$$

$$|\tilde{c}| \leq \|\mathbf{L}(\mathbf{A}_1 - \alpha\mathbf{A}_2)\mathbf{L}^T\|\|\mathbf{w}_k\|^2 + 2\|\mathbf{b}_1 - \alpha\mathbf{b}_2\|\|\mathbf{L}\|\|\mathbf{w}_k\| + |c_1 - \alpha c_2|, \tag{A.8}$$

where for a given matrix $\mathbf{M}$, $\|\mathbf{M}\|$ denotes the spectral norm $\|\mathbf{M}\| = \sqrt{\lambda_{\max}(\mathbf{M}^T\mathbf{M})}$. Using (A.4) along with (A.7) and (A.8) we have $\|\tilde{\mathbf{b}}\| \leq U_1$ and $|\tilde{c}| \leq U_2$. Therefore,

$$\|\mathbf{v}_k\| \leq 2\sqrt{\frac{U_1^2}{\lambda_{\min}(\widetilde{\mathbf{A}})^2} + \frac{U_2}{\lambda_{\min}(\widetilde{\mathbf{A}})}}, \tag{A.9}$$

and hence,

$$\|\mathbf{x}_k\| \overset{(A.3)}{=} \|\mathbf{F}\mathbf{v}_k + \mathbf{L}^T\mathbf{w}_k\| \leq \|\mathbf{F}\mathbf{v}_k\| + \|\mathbf{L}^T\mathbf{w}_k\| \leq \|\mathbf{v}_k\| + \|\mathbf{L}\|\|\mathbf{w}_k\|$$

$$\overset{(A.4),(A.9)}{\leq} 2\sqrt{\frac{U_1^2}{\lambda_{\min}(\widetilde{\mathbf{A}})^2} + \frac{U_2}{\lambda_{\min}(\widetilde{\mathbf{A}})}} + \|\mathbf{L}\|\frac{\sqrt{\rho}}{\lambda_{\min}(\mathbf{L}\mathbf{L}^T)}.$$

$\square$

References

1. Anstreicher, K., Chen, X., Wolkowicz, H., Yuan, Y.: Strong duality for a trust-region type relaxation of the quadratic assignment problem. Linear Algebra Appl. **301**(1–3), 121–136 (1999)
2. Anstreicher, K., Wolkowicz, H.: On Lagrangian relaxation of quadratic matrix constraints. SIAM J. Matrix Anal. Appl. **22**(1), 41–55 (2000)
3. Auslender, A., Teboulle, M.: Asymptotic Cones and Functions in Optimization and Variational Inequalities. Springer, Heidelberg (2003)
4. Beck, A.: Quadratic matrix programming. SIAM J. Optim. **17**(4), 1224–1238 (2006)
5. Beck, A., Ben-Tal, A., Teboulle, M.: Finding a global optimal solution for a quadratically constrained fractional quadratic problem with applications to the regularized total least squares. SIAM J. Matrix Anal. Appl. **28**(2), 425–445 (2006)
6. Beck, A., Eldar, Y.C.: Strong duality in nonconvex quadratic optimization with two quadratic constraints. SIAM J. Optim. **17**(3), 844–860 (2006)
7. Ben-Tal, A., Nemirovski, A.: Lectures on Modern Convex Optimization. MPS-SIAM Series on Optimization (2001)
8. Ben-Tal, A., Teboulle, M.: Hidden convexity in some nonconvex quadratically constrained quadratic programming. Math. Program. **72**(1), 51–63 (1996)
9. Dinkelbach, W.: On nonlinear fractional programming. Math. Program. **13**, 492–498 (1996)
10. Fortin, C., Wolkowicz, H.: The trust region subproblem and semidefinite programming. Optim. Methods Softw. **19**(1), 41–67 (2004)
11. Golub, G.H., Hansen, P.C., O'Leary, D.P.: Tikhonov regularization and total least squares. SIAM J. Matrix Anal. Appl. **21**(2), 185–194 (1999)
12. Golub, G.H., Van Loan, C.F.: An analysis of the total least-squares problem. SIAM J. Numer. Anal. **17**(6), 883–893 (1980)
13. Guo, H., Renaut, R.: A regularized total least squares algorithm. In:Total Least Squares and Errors-in-Variables Modeling, pp. 57–66. Kluwer, Dordrecht (2002)
14. Van Huffel, S., Vandewalle, J.: The total least-squares problem: computational aspects and analysis, vol. 9 of Frontier in applied mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1991)
15. Moré, J.J.: Generalization of the trust region problem. Optim. Methods Softw. **2**, 189–209 (1993)

16. Nesterov, Y., Nemirovskii, A.: Interior-point polynomial algorithms in convex programming, vol. 13 of SIAM Studies in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1994)
17. Polyak, B.T.: Convexity of quadratic transformations and its use in control and optimization. J. Optim. Theory Appl. **99**(3), 553–583 (1998)
18. Renaut, R.A., Guo, H.: Efficient algorithms for solution of regularized total least squares. SIAM J. Matrix Anal. Appl. **26**(2), 457–476 (2005)
19. Rendl, F., Wolkowicz, H.: A semidefinite framework for trust region subproblems with applications to large scale minimization. Math. Program. **77**(2, Ser.B), 273–299 (1997)
20. Sima, D., Van Huffel, S., Golub, G.H.: Regularized total least squares based on quadratic eigenvalue problem solvers. BIT Numer. Math. **44**(4), 793–812 (2004)
21. Sorensen, D.C.: Minimization of a large-scale quadratic function subject to a spherical constraint. SIAM J. Optim. **7**(1), 141–161 (1997)
22. Stern, R.J., Wolkowicz, H.: Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. SIAM J. Optim. **5**(2), 286–313 (1995)
23. Ye, Y., Zhang, S.: New results on quadratic minimization. SIAM J. Optim. **14**, 245–267 (2003)