A Dual Optimization Approach to Inverse Quadratic Eigenvalue Problems with Partial Eigenstructure

Zheng-Jian Bai [*] Delin Chu [†] Defeng Sun [‡]

June 1, 2005

Abstract

The inverse quadratic eigenvalue problem (IQEP) arises in the field of structural dynamics. It aims to find three symmetric matrices, known as the mass, the damping and the stiffness matrices, respectively such that they are closest to the given analytical matrices and satisfy the measured data. The difficulty of this problem lies in the fact that in applications the mass matrix should be positive definite and the stiffness matrix positive semidefinite. Based on an equivalent dual optimization version of the IQEP, we present a quadratically convergent Newton-type method. Our numerical experiments confirm the high efficiency of the proposed method.

1 Introduction

Throughout this paper the following notation will be used:

- $A \succeq 0$ ($A \succ 0$, respectively) means that the $n \times n$ real matrix A is symmetric and positive semidefinite (symmetric and positive definite, respectively).
- A^T, $\sigma(A)$, and $\|A\|$ denote the transpose, the spectrum, and the Frobenius norm of $A \in \mathbb{R}^{n\times n}$, respectively.
- I: the identity matrix of an appropriate dimension.
- $\mathcal{S}^n$ and $\mathcal{S}^n_+$ denote the set of all real $n \times n$ symmetric matrices and the cone of positive semidefinite matrices in this set, respectively. Let

$$\Omega_0 := \mathcal{S}^n \times \mathcal{S}^n \times \mathcal{S}^n \quad \text{and} \quad \Omega := \{(M, C, K) \in \Omega_0 \,:\, M \succeq 0, \quad K \succeq 0\}.$$

[*](matbzj@nus.edu.sg) Department of Mathematics, National University of Singapore, 2 Science Drive 2, Republic of Singapore.

[†](matchudl@nus.edu.sg) Department of Mathematics, National University of Singapore, 2 Science Drive 2, Republic of Singapore.

[‡](matsundf@nus.edu.sg) Department of Mathematics, National University of Singapore, 2 Science Drive 2, Republic of Singapore. This author's research was partially supported by Grant R146-000-061-112 of the National University of Singapore. Fax: (65) 6779 5452.

For given matrices $M, C, K \in \mathbb{R}^{n\times n}$, let

$$Q(\lambda) := \lambda^2 M + \lambda C + K.$$

Then the quadratic eigenvalue problem (QEP) is to find scalars $\lambda \in \mathbb{C}$ and nonzero vectors $\mathbf{x}$ such that

$$Q(\lambda)\mathbf{x} = 0, \tag{1}$$

where λ and $\mathbf{x}$ are called the eigenvalue and the eigenvector, respectively. For various applications, mathematical properties, and numerical solution techniques of the QEP, we refer the readers to the survey paper [44]. In many practical applications, the matrices M, C, and K are required to be symmetric. In particular, in one of the most popular numerical models – the finite element model (FEM) [18, 44]:

$$M\ddot{x}(t) + C\dot{x}(x) + Kx(t) = 0, \tag{2}$$

where the mass matrix M, the damping matrix C, and the stiffness matrix K are required to be symmetric with M and K being positive definite and positive semidefinite, respectively. It is shown in [44] that the general solution to the homogeneous equation (2) is given in terms of the solution of the QEP (1). Unfortunately, the natural frequencies and model shapes (eigenvalues and eigenvectors) predicted by the FEM often disagree with the physical experimental results. The FEM updating aims to improve the numerical model (2) by the measured data. Various model updating methods have been presented in the literature [18, 26, 28, 32]. Mathematically, the model updating problem is an inverse quadratic eigenvalue problem (IQEP) which aims to find matrices M, C, and K such that they are closest to the given estimated analytical matrices and satisfy the partially measured eigen-data. The general IQEP can be defined as follows:

- Given a measured partial eigenpair $(\Lambda, X) \in \mathbb{R}^{k\times k} \times \mathbb{R}^{n\times k}$ with

$$1 \le k \le n, \quad \text{rank}(X) = k, \quad \text{and} \quad \Lambda = \text{diag}\{\Lambda_1, \ldots, \Lambda_\mu, \Lambda_{\mu+1}, \ldots, \Lambda_\nu\}, \tag{3}$$

  where

$$\Lambda_i = \text{diag}\{\overbrace{\lambda_i^{[2]}, \ldots, \lambda_i^{[2]}}^{s_i}\} \quad \text{for} \quad 1 \le i \le \mu \quad \text{and} \quad \Lambda_i = \lambda_i I_{s_i} \quad \text{for} \quad \mu+1 \le i \le \nu,$$

$$\lambda_i^{[2]} = \begin{bmatrix} \alpha_i & \beta_i \\ -\beta_i & \alpha_i \end{bmatrix} \in \mathbb{R}^{2\times 2}, \quad \beta_i \neq 0, \quad \sigma(\Lambda_i) \cap \sigma(\Lambda_j) = \emptyset, \quad \forall\, 1 \le i \neq j \le \mu,$$

  and

$$\lambda_i \in \mathbb{R}, \quad \lambda_i \neq \lambda_j, \quad \forall\, \mu+1 \le i \neq j \le \nu,$$

  find $M, C, K \in \mathcal{S}^n$ with $M \succ 0$ and $K \succeq 0$ such that

$$MX\Lambda^2 + CX\Lambda + KX = 0. \tag{4}$$

The model updating problem, for given $M_a, C_a, K_a \in \mathcal{S}^n$, which are called the estimated analytic mass, damping, and stiffness matrices, respectively in the FEM updating [18], is then to find $M, C, K \in \mathcal{S}^n$ such that (M, C, K) is a solution to the following optimization problem

$$\begin{array}{ll} \inf & \frac{c_1}{2}\|M - M_a\|^2 + \frac{c_2}{2}\|C - C_a\|^2 + \frac{1}{2}\|K - K_a\|^2 \\ \text{s.t.} & MX\Lambda^2 + CX\Lambda + KX = 0, \\ & M \succ 0, \quad C = C^T, \quad K \succeq 0, \end{array} \tag{5}$$

where c_1 and c_2 are two positive weighting parameters. As in [24], let the QR factorization of X be given by

$$X = Q \begin{bmatrix} R \\ 0 \end{bmatrix}, \tag{6}$$

where $Q \in \mathbb{R}^{n\times n}$ is orthogonal and $R \in \mathbb{R}^{k\times k}$ is nonsingular and upper triangular. By renaming $M := \sqrt{c_1}Q^T M Q$, $C := \sqrt{c_2}Q^T C Q$, $K := Q^T K Q$, $M_a := \sqrt{c_1}Q^T M_a Q$, $C_a := \sqrt{c_2}Q^T C_a Q$, and $K_a := Q^T K_a Q$, without loss of generality, we may assume from now on that problem (5) takes the following form

$$\begin{array}{ll} \inf & \frac{1}{2}\|M - M_a\|^2 + \frac{1}{2}\|C - C_a\|^2 + \frac{1}{2}\|K - K_a\|^2 \\ \text{s.t.} & \frac{1}{\sqrt{c_1}}M \begin{bmatrix} R \\ 0 \end{bmatrix} \Lambda^2 + \frac{1}{\sqrt{c_2}}C \begin{bmatrix} R \\ 0 \end{bmatrix} \Lambda + K \begin{bmatrix} R \\ 0 \end{bmatrix} = 0, \\ & M \succ 0, \quad C = C^T, \quad K \succeq 0. \end{array} \tag{7}$$

Theorem 1.1 *Problem (7) admits a strictly feasible solution if and only if Λ is nonsingular.*

Proof: First we show the *necessity* by contradiction. Assume that Λ is singular. Then there exists $y \in \mathbb{R}^k$ with $y \neq 0$ such that

$$\Lambda y = 0.$$

As a result, for any feasible solution (M, C, K) to problem (7), we have

$$\left(\frac{1}{\sqrt{c_1}}M \begin{bmatrix} R \\ 0 \end{bmatrix} \Lambda^2 + \frac{1}{\sqrt{c_2}}C \begin{bmatrix} R \\ 0 \end{bmatrix} \Lambda + K \begin{bmatrix} R \\ 0 \end{bmatrix}\right) y = 0, \quad \text{i.e.,} \quad K \begin{bmatrix} Ry \\ 0 \end{bmatrix} = 0.$$

Note that R is nonsingular, which yields that $Ry \neq 0$. Hence K must be singular. This is a contraction with the existence of a strictly feasible solution to problem (7). Therefore, problem (7) can admit a strictly feasible solution only if Λ is nonsingular.

Next we show the *sufficiency*. Following the idea of [11], we can obtain a special feasible solution to (7) as follows:

$$\overline{M} = \sqrt{c_1}\begin{bmatrix} (RR^T)^{-1} & 0 \\ 0 & I \end{bmatrix}, \ \overline{C} = \sqrt{c_2}\begin{bmatrix} -R^{-T}(\Lambda + \Lambda^T)R^{-1} & 0 \\ 0 & 0 \end{bmatrix}, \ \overline{K} = \begin{bmatrix} R^{-T}\Lambda^T\Lambda R^{-1} & 0 \\ 0 & I \end{bmatrix}. \tag{8}$$

Since both R and Λ are nonsingular, $\overline{M}$ and $\overline{K}$ are symmetric and positive definite. This implies that $(\overline{M}, \overline{C}, \overline{K})$ above is a strictly feasible solution to problem (7). $\square$

Because of Theorem 1.1 and the generalized Slater condition (13) (see Section 2) used in our dual approach, in the following development we make a blanket assumption, unless stated otherwise, that the following condition holds:

Assumption A.1 Λ is nonsingular.

Remark 1.2 *Problem (7) fails to possess a strictly feasible solution if Λ is singular. As a remedy, we can reduce problem (7) equivalently to a similar problem which admits a strictly feasible solution. See the last part of Section 4 for discussions on this.*

It is noted that in numerical computation the condition $M \succ 0$ is better to be replaced by $M \succeq 0$. This consideration leads us to focus on the following optimization problem

$$\begin{array}{ll} \min & \frac{1}{2}\|M - M_a\|^2 + \frac{1}{2}\|C - C_a\|^2 + \frac{1}{2}\|K - K_a\|^2 \\ \text{s.t.} & \frac{1}{\sqrt{c_1}} M \begin{bmatrix} R \\ 0 \end{bmatrix} \Lambda^2 + \frac{1}{\sqrt{c_2}} C \begin{bmatrix} R \\ 0 \end{bmatrix} \Lambda + K \begin{bmatrix} R \\ 0 \end{bmatrix} = 0, \\ & (M, C, K) \in \Omega. \end{array} \tag{9}$$

Actually, let $(M^*, C^*, K^*) \in \Omega$ be the optimal solution to problem (9). Then for any $\varepsilon > 0$,

$$(M^* + \varepsilon\overline{M}, C^* + \varepsilon\overline{C}, K^* + \varepsilon\overline{K})$$

is an $O(\varepsilon)$-optimal solution to problem (7), where $(\overline{M}, \overline{C}, \overline{K})$ is defined by (8). So, without causing any confusion, we refer the above optimization problem as the IQEP.

Due to complicated structures of the IQEP and lack of tools dealing with it, many authors considered various simplified versions of the IQEP by dropping or relaxing the requirement on the positive semidefiniteness of matrices M and K [2, 3, 4, 5, 13, 17, 18, 25, 47, 48, 49]. As observed in the literature of structural dynamics [17, 18], however, these approaches may fail to guarantee the positive semidefiniteness of M and K, which is vital in applications.

In this paper, we directly deal with the IQEP with the required positive semidefiniteness condition on matrices M and K. We shall introduce a generalized Newton method to the dual of problem (9). This approach is motivated by two recent papers due to Chu, Kuo, and Lin [11] and Qi and Sun [35]. In [11], Chu, Kuo, and Lin made an important step by showing for the first time that the general IQEP admits a nontrivial solution, i.e., there exist symmetric matrices M, C, and K with $M \succ 0$ and $K \succeq 0$ satisfying (4). In [35], based on recent developments on strongly semismooth matrix valued functions [43], Qi and Sun presented a highly efficient quadratically convergent Newton method for solving the nearest correlation matrix problem proposed by Higham in [21] and further studied by Malick [29], Boyd and Lin [9], and Toh, Tütüncü, and Todd [46]. In order to prove that our proposed Newton method possesses a quadratic convergence rate, we need the generalized Jacobian used in our method to be positive definite at a solution point. We shall address this important question under a general setting in this paper.

The organization of this paper is as follows. In Section 2, we discuss some basic convex optimization theory used in this paper and present some results regarding Clarke's generalized Jacobian of certain locally Lipschitz functions. In Section 3, we present the dual forms of the IQEP. In Section 4, we give the details of our generalized Newton method and its convergence analysis. We report our numerical results in Section 5 and make some final conclusions in Section 6.

2 Preliminaries

Consider the following convex optimization problem

$$\begin{array}{rll} f^* := & \inf & f(x) \\ & \text{s.t.} & \mathcal{A}x = b, \\ & & x \in \mathcal{Q}, \end{array} \tag{10}$$

where $f : \mathcal{X} \to \mathbb{R}$ is a convex function, $A : \mathcal{X} \to \mathcal{Y}$ is a linear operator, $b \in \mathcal{Y}$, $\mathcal{Q}$ is a closed convex cone in $\mathcal{X}$, and $\mathcal{X}$ and $\mathcal{Y}$ are finite dimensional real vector spaces each equipped with a scalar inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. Let $\mathcal{A}^* : \mathcal{Y} \to \mathcal{X}$ be the adjoint of $\mathcal{A}$. Define the Lagrangian function $\mathcal{K} : \mathcal{Q} \times \mathcal{Y} \to \mathbb{R}$ by

$$\mathcal{K}(x, y) := f(x) + \langle y, b - \mathcal{A}x \rangle, \quad (x, y) \in \mathcal{Q} \times \mathcal{Y}. \tag{11}$$

Let $g : \mathcal{Y} \to \mathbb{R}$ be defined by

$$g(y) := \inf_{x \in \mathcal{Q}} \mathcal{K}(x, y), \quad y \in \mathcal{Y}.$$

Then the dual of problem (10) takes the following form:

$$\begin{array}{ll} \sup & g(y) \\ \text{s.t.} & y \in \mathcal{Y}. \end{array} \tag{12}$$

Recall that the generalized Slater condition is said to hold for the convex optimization problem (10) if

$$\begin{cases} \mathcal{A} : \mathcal{X} \to \mathcal{Y} \text{ is onto,} \\ \exists\ \bar{x} \in \mathcal{X} \text{ such that } \ \mathcal{A}\bar{x} = b,\ \bar{x} \in \operatorname{int}(\mathcal{Q}), \end{cases} \tag{13}$$

where "int" denotes the topological interior of a given set. The classical duality theory for convex programming [40, Theorems 17 & 18] says that under the generalized Slater condition (13), the following hold:

- There exists at least one y^* such that

$$g(y^*) = \sup_{y \in \mathcal{Y}} g(y) = \inf_{x \in \mathcal{Q}} \sup_{y \in \mathcal{Y}} \mathcal{K}(x, y) = f^*$$

  and if x^* uniquely solves the following problem

$$\begin{array}{ll} \inf & \mathcal{K}(x, y^*) \\ \text{s.t.} & x \in \mathcal{Q}, \end{array} \tag{14}$$

  then x^* solves the original problem (10).

- For every real number τ, the set $\{y \in \mathcal{Y} : g(y) \geq \tau\}$ is closed, bounded, and convex.

Let x^0 be a given vector in $\mathcal{X}$. When $f : \mathcal{X} \to \mathbb{R}$ is given by the following quadratic function

$$f(x) := \frac{1}{2}\langle x - x^0, x - x^0\rangle, \quad x \in \mathcal{X},$$

the convex optimization problem (10) becomes

$$\begin{array}{ll} \min & \frac{1}{2}\langle x - x^0, x - x^0\rangle \\ \text{s.t.} & \mathcal{A}x = b, \\ & x \in \mathcal{Q}, \end{array} \tag{15}$$

which in the literature is known as the best approximation in Hilbert space. For a survey on the latter, see [34]. Here, we use "min" instead of "inf" as in this case the infimum is always attainable. Let $\mathcal{D} \subseteq \mathcal{X}$ be a closed convex set. For any $x \in \mathcal{X}$, let $\Pi_{\mathcal{D}}(x)$ denote the metric projection of x onto $\mathcal{D}$, i.e., $\Pi_{\mathcal{D}}(x)$ is the unique optimal solution to the following convex quadratic programming problem

$$\begin{array}{ll} \min & \frac{1}{2}\langle z - x, z - x\rangle \\ \text{s.t.} & z \in \mathcal{D}. \end{array} \tag{16}$$

Then, the dual of problem (15) has the following form:

$$\begin{array}{ll} \sup & -\theta(y) \\ \text{s.t.} & y \in \mathcal{Y}, \end{array} \tag{17}$$

where $\theta : \mathcal{Y} \to \mathbb{R}$ is given by

$$\theta(y) := \frac{1}{2}\|x^0 + \mathcal{A}^*y\|^2 - \frac{1}{2}\|x^0 + \mathcal{A}^*y - \Pi_{\mathcal{Q}}(x^0 + \mathcal{A}^*y)\|^2 - \langle b, y\rangle - \frac{1}{2}||x^0\|^2. \tag{18}$$

In particular, if $\mathcal{Q}$ is a closed convex cone, by [50], we have

$$\theta(y) = \frac{1}{2}\|\Pi_{\mathcal{Q}}(x^0 + \mathcal{A}^*y)\|^2 - \langle b, y\rangle - \frac{1}{2}||x^0\|^2. \tag{19}$$

From [50] we know that θ is a continuously differentiable convex function and its gradient at $y \in \mathcal{Y}$ is given by

$$\nabla\theta(y) = \mathcal{A}\,\Pi_{\mathcal{Q}}(x^0 + \mathcal{A}^*y) - b. \tag{20}$$

Furthermore, since $\Pi_{\mathcal{Q}}(\cdot)$ is globally Lipschitz continuous with modulus 1, $\nabla\theta(\cdot)$ is globally Lipschitz continuous. Additionally, if the Slater condition (13) holds, then for any $\tau \in \mathbb{R}$, the level set $\{y \in \mathcal{Y} : \theta(y) \leq \tau\}$ is bounded. Therefore, one may use any gradient based method like the steepest decent method or quasi-Newton methods to find an optimal solution y^* to (17) first and then solve (14) to get an optimal solution to problem (15). This falls within the dual approach outlined by Rockafellar in [40, Page 4] for general convex optimization problems. Note that in this case the unique solution to (14) is given by

$$x^* = \Pi_{\mathcal{Q}}(x^0 + \mathcal{A}^*y^*). \tag{21}$$

Since in general $\Pi_{\mathcal{Q}}(\cdot)$ is not continuously differentiable, $\theta(\cdot)$ is not twice continuously differentiable. So we cannot directly use Newton's method to solve (17), i.e.,

$$\begin{array}{ll} \min & \theta(y) \\ \text{s.t.} & y \in \mathcal{Y}. \end{array} \tag{22}$$

Fortunately, we can apply Clarke's Jacobian based Newton method developed in the last two decades for solving locally Lipschitz equations [23, 36].

Let $\mathcal{Z}$ be an arbitrary finite dimensional real vector space. Let $\mathcal{O}$ be an open set in $\mathcal{Y}$ and $\Xi : \mathcal{O} \subseteq \mathcal{Y} \to \mathcal{Z}$ be a locally Lipschitz continuous function on the open set $\mathcal{O}$. Rademacher's theorem [41, Chapter 9.J] says that Ξ is almost everywhere Fréchet differentiable in $\mathcal{O}$. We denote by $\mathcal{O}_\Xi$ the set of points in $\mathcal{O}$ where Ξ is Fréchet differentiable. Let $\Xi'(y)$ denote the Jacobian of Ξ at $y \in \mathcal{O}_\Xi$. Then Clarke's generalized Jacobian of Ξ at $y \in \mathcal{O}$ is defined by [12]

$$\partial \Xi(y) := \mathrm{conv}\{\partial_B \Xi(y)\},$$

where "conv" denotes the convex hull and

$$\partial_B \Xi(y) := \left\{ V \,:\, V = \lim_{j \to \infty} \Xi'(y^j)\,,\; y^j \in \mathcal{O}_\Xi \right\}.$$

Suppose that $F : \mathcal{O} \subseteq \mathcal{Y} \to \mathcal{Y}$ is a locally Lipschitz continuous function on the open set $\mathcal{O}$. Let y^0 be an arbitrary point in $\mathcal{O}$. Then Clarke's Jacobian based Newton method for solving the locally Lipschitz equation $F(y) = 0$ can be described by

$$y^{j+1} = y^j - V_j^{-1} F(y^j), \qquad V_j \in \partial F(y^j), \quad j = 0, 1, 2, \ldots. \tag{23}$$

In general, the above iterative method does not converge even if $\{\|V_j^{-1}\|\}$ is uniformly bounded and y^0 is arbitrarily close to a given solution of $F(y) = 0$ [23]. Kummer in [23] proposed a general condition guaranteeing the superlinear convergence of (23), which generalized Kojima and Shindo's condition for superlinear (quadratic) convergence of Newton's method for piecewise smooth equations [22]. Kummer's result was largely unnoticed until Qi and Sun in [36] published their now well known work by showing that the iteration sequence generated by (23) can converge superlinearly if F is a *semismooth function*.

The concept of semismoothness was introduced by Mifflin [31] for functionals. In order to study the convergence of the iterative method (23), Qi and Sun [36] extended the definition of semismoothness to vector-valued functions. The following definition of semismoothness is convenient in use.

Definition 2.1 *Let $\Xi : \mathcal{O} \subseteq \mathcal{Y} \to \mathcal{Z}$ be a locally Lipschitz continuous function on the open set $\mathcal{O}$. We say that Ξ is semismooth at a point $y \in \mathcal{O}$ if*

(i) *Ξ is directionally differentiable at y; and*

(ii) *for any $x \to y$ and $V \in \partial \Xi(x)$,*

$$\Xi(x) - \Xi(y) - V(x - y) = o(||x - y||)\,. \tag{24}$$

A stronger notion than semismoothness is *strong semismoothness*. A locally Lipschitz continuous function $\Xi : \mathcal{O} \subseteq \mathcal{Y} \to \mathcal{Z}$ is said to be strongly semismooth at a point $y \in \mathcal{O}$ if F is semismooth at y and for any $x \to y$ and $V \in \partial\Xi(x)$,

$$\Xi(x) - \Xi(y) - V(x-y) = O(\|x-y\|^2)\,. \tag{25}$$

The following convergence result is established in [36, Theorem 3.2] while the result on superlinear convergence can also be found in [23, Proposition 3].

Proposition 2.2 *Let $F : \mathcal{O} \subseteq \mathcal{Y} \to \mathcal{Y}$ be a locally Lipschitz continuous function on the open set $\mathcal{O}$. Let $y^* \in \mathcal{O}$ be such that $F(y^*) = 0$. Assume that F is semismooth at y^* and that every element in $\partial F(y^*)$ is nonsingular. Then every sequence generated by (23) converges to y^* superlinearly provided that the starting point y^0 is sufficiently close to y^*. Moreover, if F is strongly semismooth at y^*, the rate of convergence is quadratic.*

Now, we consider the following equation:

$$F(y) := \nabla\theta(y) = \mathcal{A}\,\Pi_{\mathcal{Q}}(x^0 + \mathcal{A}^* y) - b = 0, \quad y \in \mathcal{Y}. \tag{26}$$

As we have already known that F is a globally Lipschitz continuous mapping and so $\partial F(y)$ is well defined everywhere in $\mathcal{Y}$. In order to apply Clarke's Jacobian based Newton method (23) to find a root of (26), we need to assume that the two conditions imposed in Proposition 2.2 are valid. Certainly, the (strong) semismoothness of F at y^* follows easily if the metric projector $\Pi_{\mathcal{Q}}(\cdot)$ is (strongly) semismooth at $x^0 + \mathcal{A}^* y^*$. This means that if one knows in advance that $\Pi_{\mathcal{Q}}(\cdot)$ is (strongly) semismooth everywhere in $\mathcal{Y}$ (it is the case for the IQEP discussed in this paper), then there is no need to be concerned about the semismoothness assumption on F. However, the second condition, i.e., the nonsingularity assumption on the generalized Jacobian $\partial F(y^*)$, needs more elaborations. We shall discuss this issue in the remaining part of this section.

Let $F : \mathcal{Y} \to \mathcal{Y}$ be defined by (26). The next proposition on Clarke's Jacobian of F follows easily from [12, Page 75].

Proposition 2.3 *Let $\bar{y} \in \mathcal{Y}$ and $\bar{x} := x^0 + \mathcal{A}^*\bar{y}$. Then, for any $d \in \mathcal{Y}$, it holds that*

$$\partial F(\bar{y})d \subseteq \mathcal{A}\,\partial\Pi_{\mathcal{Q}}(\bar{x})\mathcal{A}^* d.$$

Proposition 2.3 relates the Jacobian of F to the Jacobian of the metric projector $\Pi_{\mathcal{Q}}(\cdot)$, for which we have the following result from [30, Proposition 1].

Lemma 2.4 *Let $\mathcal{D} \subseteq \mathcal{X}$ be a closed convex set. Then, for any $x \in \mathcal{X}$ and $V \in \partial\Pi_{\mathcal{D}}(x)$,*

(i) *V is self-adjoint.*

(ii) $\langle d, Vd\rangle \geq 0 \quad \forall\, d \in \mathcal{X}$.

(iii) $\langle Vd, d - Vd\rangle \geq 0 \quad \forall\, d \in \mathcal{X}$.

Let $\mathcal{D} \subseteq \mathcal{X}$ be a closed convex set. For any $x \in \mathcal{X}$, denote

$$\mathrm{dist}(x, \mathcal{D}) := \inf\{||x - d|| \, : \, d \in \mathcal{D}\}.$$

The *tangent cone* of $\mathcal{D}$ at a point $x \in \mathcal{D}$, which we denote $\mathcal{T}_{\mathcal{D}}(x)$, consisting of all tangent vectors of $\mathcal{D}$ at x, is defined by

$$\mathcal{T}_{\mathcal{D}}(x) := \{d \in \mathcal{X} \, : \, \mathrm{dist}(x + td, \mathcal{D}) = o(t) \, , \, t \geq 0\}. \tag{27}$$

For any $x \in \mathcal{D}$, we denote by $\mathrm{lin}\,(\mathcal{T}_{\mathcal{D}}(x))$ the lineality space of $\mathcal{T}_{\mathcal{D}}(x)$, i.e., the largest linear space in $\mathcal{T}_{\mathcal{D}}(x)$.

The following definition of constraint nondegeneracy is adapted from the works of Robinson, who defined it for general optimization problems and variational inequalities [37, 38, 39].

Definition 2.5 *We say that a feasible point x to the convex optimization problem (10) is constraint nondegenerate if*

$$\begin{cases} \mathcal{A}\mathcal{X} = \mathcal{Y}, \\ \mathcal{X} + \mathrm{lin}\,(\mathcal{T}_{\mathcal{Q}}(x)) = \mathcal{X} \end{cases} \tag{28}$$

or equivalently if

$$\mathcal{A}\,\mathrm{lin}\,(\mathcal{T}_{\mathcal{Q}}(x)) = \mathcal{Y}. \tag{29}$$

The constraint nondegenerate condition (28) in its general form is extensively used by Bonnans and Shapiro in [8] for perturbation analysis of optimization problems. In the context of semidefinite programming, condition (28) reduces to the primal nondegenerate condition proposed by Alizadeh, Haeberly, and Overton [1]. For the conventional nonlinear programming problem with finitely many equality and inequality constraints, the constraint nondegenerate condition is equivalent to the well known linear independence constraint qualification [37, 42].

Definition 2.6 *Let $\mathcal{D} \subseteq \mathcal{X}$ be a closed convex set. We say that the metric projector $\Pi_{\mathcal{D}}(\cdot)$ is Jacobian amicable at $x \in \mathcal{X}$ if for any $V \in \partial\Pi_{\mathcal{D}}(x)$ and $d \in \mathcal{X}$ such that $Vd = 0$, it holds that*

$$d \in (\mathrm{lin}\,(\mathcal{T}_{\mathcal{D}}(x_+)))^{\perp},$$

where $x_+ := \Pi_{\mathcal{D}}(x)$ and $(\mathrm{lin}\,(\mathcal{T}_{\mathcal{D}}(x_+)))^{\perp}$ is defined by

$$(\mathrm{lin}\,(\mathcal{T}_{\mathcal{D}}(x_+)))^{\perp} := \{d \in \mathcal{X} \, : \, \langle d, h\rangle = 0 \quad \forall \; h \in \mathrm{lin}\,(\mathcal{T}_{\mathcal{D}}(x_+))\}.$$

We say that $\Pi_{\mathcal{D}}(\cdot)$ is Jacobian amicable if it is Jacobian amicable at every point in $\mathcal{X}$.

Now, we are ready to state our result on the nonsingularity of Clarke's Jacobian of the mapping F defined by (26).

Proposition 2.7 *Let $\bar{y} \in \mathcal{Y}$ be such that $F(\bar{y}) = b$. Let $\bar{x} := x^0 + \mathcal{A}^*\bar{y}$ and $\bar{x}_+ := \Pi_{\mathcal{Q}}(\bar{x})$. Assume that the constraint nondegenerate condition (29) holds at $\bar{x}_+$ and that $\Pi_{\mathcal{Q}}(\cdot)$ is Jacobian amicable at $\bar{x}$. Then every element in $\mathcal{A}\partial\Pi_{\mathcal{Q}}(\bar{x})\mathcal{A}^*$ is self-adjoint and positive definite.*

Proof: Let V be an arbitrary element in $\mathcal{A}\partial\Pi_{\mathcal{Q}}(\bar{x})\mathcal{A}^*$. Then, there exists an element $W \in \partial\Pi_{\mathcal{Q}}(\bar{x})$ such that

$$V = \mathcal{A}W\mathcal{A}^*.$$

Since, by Lemma 2.4, W is self-adjoint and positive semidefinite, we know that V is also self-adjoint and positive semidefinite.

Next, we show the positive definiteness of V. Let $d \in \mathcal{X}$ be such that $Vd = 0$. Then, by (iii) of Lemma 2.4, we obtain that

$$0 = \langle d, Vd\rangle = \langle d, \mathcal{A}W\mathcal{A}^*d\rangle = \langle \mathcal{A}^*d, W\mathcal{A}^*d\rangle \ \geq \langle W\mathcal{A}^*d, W\mathcal{A}^*d\rangle,$$

which implies

$$W\mathcal{A}^*d = 0.$$

Therefore, by the assumption that $\Pi_{\mathcal{Q}}(\cdot)$ is Jacobian amicable at $\bar{x}$, we know that

$$\mathcal{A}^*d \in \left(\mathrm{lin}\left(\mathcal{T}_{\mathcal{Q}}(\bar{x}_+)\right)\right)^{\perp}.$$

Since the constraint nondegenerate condition (29) holds at $\bar{x}_+$, there exists a vector $h \in \mathrm{lin}\left(\mathcal{T}_{\mathcal{Q}}(\bar{x}_+)\right)$ such that $\mathcal{A}h = d$. Hence, since $\mathcal{A}^*d \in \left(\mathrm{lin}\left(\mathcal{T}_{\mathcal{Q}}(\bar{x}_+)\right)\right)^{\perp}$ and $h \in \mathrm{lin}\left(\mathcal{T}_{\mathcal{Q}}(\bar{x}_+)\right)$, it holds that

$$\langle d, d\rangle = \langle d, \mathcal{A}h\rangle = \langle \mathcal{A}^*d, h\rangle = 0.$$

Thus $d = 0$. This, together with the fact that V is self-adjoint and positive semidefinite, shows that V is self-adjoint and positive definite. □

Proposition 2.7 motivates us to define the following modified version of the Jacobian based Newton method (23) to solve equation (26):

$$y^{j+1} = y^j - V_j^{-1}F(y^j), \quad V_j := \mathcal{A}W_j\mathcal{A}^*,\ W_j \in \partial\Pi_{\mathcal{Q}}(x^0 + \mathcal{A}^*y^j), \quad j = 0, 1, 2, \ldots, \tag{30}$$

where $y^0 \in \mathcal{Y}$ is an initial point.

By Proposition 2.7, we can obtain the following convergence results for the iterative method (30) in a similar way as Proposition 2.2.

Proposition 2.8 *Let $\bar{y} \in \mathcal{Y}$ be such that $F(\bar{y}) = b$. Let $\bar{x} := x^0 + \mathcal{A}^*\bar{y}$ and $\bar{x}_+ := \Pi_{\mathcal{Q}}(\bar{x})$. Assume that the constraint nondegenerate condition (29) holds at $\bar{x}_+$, $\Pi_{\mathcal{Q}}(\cdot)$ is semismooth at $\bar{x}$, and $\Pi_{\mathcal{Q}}(\cdot)$ is Jacobian amicable at $\bar{x}$. Then every sequence generated by (30) converges to $\bar{y}$ superlinearly provided that the starting point y^0 is sufficiently close to $\bar{y}$. Moreover, if $\Pi_{\mathcal{Q}}(\cdot)$ is strongly semismooth at $\bar{x}$, the rate of convergence is quadratic.*

In Proposition 2.8, besides the well studied constraint nondegenerate condition (29) at $\bar{x}_+$ in the literature, we also assume that $\Pi_{\mathcal{Q}}(\cdot)$ is Jacobian amicable at $\bar{x}$, which may hold for many interesting closed convex cones. Here we only discuss about the Jacobian amicability of $\Pi_{\mathcal{Q}}(\cdot)$ for the case that $\mathcal{Q} = \mathcal{S}^n_+$, the cone of symmetric and positive semidefinite matrices in $\mathcal{S}^n$.

Let $A \in \mathcal{S}^n$. Then A admits the following spectral decomposition

$$A = P\Sigma P^T, \tag{31}$$

where Σ is the diagonal matrix of eigenvalues of A and P is a corresponding orthogonal matrix of orthonormal eigenvectors. Under the Frobenius inner product, the projection $A_+ := \Pi_{\mathcal{S}^n_+}(A)$ of a matrix $A \in \mathcal{S}^n$ onto the cone $\mathcal{S}^n_+$ has an analytic formula

$$A_+ = P\Sigma_+ P^T,$$

where Σ_+ is the diagonal matrix whose diagonal entries are the nonnegative parts of the respective diagonal entries of Σ [20, 45]. Define three index sets of positive, zero, and negative eigenvalues of A, respectively, as

$$\alpha := \{i \,:\, \sigma_i > 0\}, \quad \beta := \{i \,:\, \sigma_i = 0\}, \quad \gamma := \{i \,:\, \sigma_i < 0\}.$$

Write

$$\Sigma = \begin{bmatrix} \Sigma_\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Sigma_\gamma \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} P_\alpha & P_\beta & P_\gamma \end{bmatrix}$$

with $P_\alpha \in \mathbb{R}^{n\times|\alpha|}$, $P_\beta \in \mathbb{R}^{n\times|\beta|}$, and $P_\gamma \in \mathbb{R}^{n\times|\gamma|}$. Define the matrix $U \in \mathcal{S}^n$ with entries

$$U_{ij} := \frac{\max\{\sigma_i, 0\} + \max\{\sigma_j, 0\}}{|\sigma_i| + |\sigma_j|}, \quad i, j = 1, \ldots, n,$$

where $0/0$ is defined to be 1. Bonnans, Cominetti and Shapiro [7] showed that $\Pi_{\mathcal{S}^n_+}(\cdot)$ is directionally differentiable everywhere in $\mathcal{S}^n$. For any $H \in \mathcal{S}^n$, Sun and Sun [43] gave the following explicit formula for $\Pi'_{\mathcal{S}^n_+}(A; H)$:

$$\Pi'_{\mathcal{S}^n_+}(A; H) = P \begin{bmatrix} P_\alpha^T H P_\alpha & P_\alpha^T H P_\beta & U_{\alpha\gamma} \circ P_\alpha^T H P_\gamma \\ P_\beta^T H P_\alpha & \Pi_{\mathcal{S}^{|\beta|}_+}\left(P_\beta^T H P_\beta\right) & 0 \\ P_\gamma^T H P_\alpha \circ U_{\alpha\gamma}^T & 0 & 0 \end{bmatrix} P^T, \tag{32}$$

where $\circ$ denotes the Hadamard product. More importantly, Sun and Sun proved that $\Pi_{\mathcal{S}^n_+}$ is a strongly semismooth matrix-valued function [43]. When A is nonsingular, i.e., $|\beta| = 0$, $\Pi_{\mathcal{S}^n_+}(\cdot)$ is continuously differentiable around A and the formula (32) reduces to the classical result of Löwner [27]. See Donoghue [15] and Bhatia [6] for detailed discussions on the latter and Chen, Qi, and Tseng [10] for a generalization.

From (32), we obtain that

$$\mathcal{T}_{\mathcal{S}^n_+}(A_+) = \{B \in \mathcal{S}^n \,:\, B = \Pi'_{\mathcal{S}^n_+}(A_+; B)\} = \{B \in \mathcal{S}^P \,:\, P_{\bar\alpha}^T B P_{\bar\alpha} \succeq 0\} \tag{33}$$

and

$$\operatorname{lin}\left(\mathcal{T}_{\mathcal{S}^n_+}(A_+)\right) = \{B \in \mathcal{S}^n \,:\, P_{\bar\alpha}^T B P_{\bar\alpha} = 0\}, \tag{34}$$

where $\bar\alpha := \{1, \ldots, n\}\backslash\alpha$ and $P_{\bar\alpha} := [P_\beta \; P_\gamma]$.

Let $\Phi(\cdot) := \Pi'_{\mathcal{S}^n_+}(A; \cdot)$. It is proved in [33, Lemma 11] that

$$\partial_B \Pi_{\mathcal{S}^n_+}(A) = \partial_B \Phi(0) \tag{35}$$

and the following operator $W : \mathcal{S}^n \to \mathcal{S}^n$ defined by

$$W(H) = P \begin{bmatrix} P_\alpha^T H P_\alpha & P_\alpha^T H P_\beta & U_{\alpha\gamma} \circ P_\alpha^T H P_\gamma \\ P_\beta^T H P_\alpha^T & 0 & 0 \\ P_\gamma^T H P_\alpha \circ U_{\alpha\gamma}^T & 0 & 0 \end{bmatrix} P^T \quad \forall\, H \in \mathcal{S}^n \tag{36}$$

is an element in $\partial_B \Pi_{\mathcal{S}_+^n}(A)$. From (32) and (35), we obtain that for any $W \in \partial \Pi_{\mathcal{S}_+^n}(A)$ and $H \in \mathcal{S}^n$ such that $WH = 0$,

$$P_\alpha^T H P_\alpha = 0, \quad P_\alpha^T H P_\beta = 0, \quad \text{and} \quad P_\alpha^T H P_\gamma = 0,$$

which, together with (34), implies

$$H \in \left(\lin \left(\mathcal{T}_{\mathcal{S}_+^n}(A_+) \right) \right)^\perp.$$

That is, $\Pi_{\mathcal{S}_+^n}(\cdot)$ is Jacobian amicable at A. For ease of reference, we summary the properties of the metric projector $\Pi_{\mathcal{S}_+^n}(\cdot)$ in the following proposition.

Proposition 2.9 *The following statements are true.*

(i) $\Pi_{\mathcal{S}_+^n}(\cdot)$ *is Fréchet differentiable at* $A \in \mathcal{S}^n$ *if and only if* A *is nonsingular.*

(ii) *For any* $A \in \mathcal{S}^n$*, the directional derivative* $\Pi'_{\mathcal{S}_+^n}(A;\cdot)$ *is Fréchet differentiable at* $H \in \mathcal{S}^n$ *if and only if* $P_\beta^T H P_\beta$ *is nonsingular.*

(iii) $\Pi_{\mathcal{S}_+^n}(\cdot)$ *is strongly semismooth everywhere in* $\mathcal{S}^n$*.*

(iv) $\Pi_{\mathcal{S}_+^n}(\cdot)$ *is Jacobian amicable everywhere in* $\mathcal{S}^n$*.*

3 Dual Reformulations

As mentioned in Section 1, we only need to focus on the IQEP under Assumption A.1, i.e., Λ is nonsingular. Let the QR factorization of X be given by (6). Denote

$$M =: \begin{bmatrix} M_1 & M_2 \\ M_2^T & M_4 \end{bmatrix},\ C := \begin{bmatrix} C_1 & C_2 \\ C_2^T & C_4 \end{bmatrix},\ K := \begin{bmatrix} K_1 & K_2 \\ K_2^T & K_4 \end{bmatrix}, \tag{37}$$

where $M_1, C_1, K_1 \in \mathcal{S}^k$, M_2, C_2, $K_2 \in \mathbb{R}^{k\times(n-k)}$, and $M_4, C_4, K_4 \in \mathcal{S}^{(n-k)}$. Let $S := R\Lambda R^{-1}$. Then problem (9) can be written equivalently as

$$\begin{array}{ll} \min & \frac{1}{2}\|M - M_a\|^2 + \frac{1}{2}\|C - C_a\|^2 + \frac{1}{2}\|K - K_a\|^2 \\ \text{s.t.} & \frac{1}{\sqrt{c_1}}(\Lambda^2)^T(R^T M_1 R) + \frac{1}{\sqrt{c_2}}\Lambda^T(R^T C_1 R) + (R^T K_1 R) = 0, \\ & \frac{1}{\sqrt{c_1}}(S^2)^T M_2 + \frac{1}{\sqrt{c_2}} S^T C_2 + K_2 = 0, \\ & (M, C, K) \in \Omega. \end{array} \tag{38}$$

We first define a linear operator $\mathcal{H} : \Omega_0 \to \mathbb{R}^{k\times k}$ by

$$\mathcal{H}(M,C,K) := \frac{1}{\sqrt{c_1}}(\Lambda^2)^T\left(R^T M_1 R\right) + \frac{1}{\sqrt{c_2}}\Lambda^T\left(R^T C_1 R\right) + \left(R^T K_1 R\right), \quad (M,C,K) \in \Omega_0. \tag{39}$$

Certainly, $\mathcal{H}$ is not necessarily surjective from Ω_0 to $\mathbb{R}^{k\times k}$. For an example, consider $\Lambda = I$. Denote

$$\mathrm{Range}(\mathcal{H}) := \{\mathcal{H}(M,C,K) \,:\, (M,C,K) \in \Omega_0\}.$$

Then $\mathrm{Range}(\mathcal{H})$ is a linear space and $\mathcal{H} : \Omega_0 \to \mathrm{Range}(\mathcal{H})$ is surjective. The adjoint $\mathcal{H}^* : \mathrm{Range}(\mathcal{H}) \to \Omega_0$ of $\mathcal{H} : \Omega_0 \to \mathrm{Range}(\mathcal{H})$ is given by

$$\mathcal{H}^*(Y) := (\mathcal{H}_1^*(Y), \mathcal{H}_2^*(Y), \mathcal{H}_3^*(Y)), \quad Y \in \mathrm{Range}(\mathcal{H}),$$

where for each $Y \in \mathrm{Range}(\mathcal{H})$,

$$\mathcal{H}_1^*(Y) = \frac{1}{2\sqrt{c_1}}\begin{bmatrix} R\Lambda^2 Y R^T + (R\Lambda^2 Y R^T)^T & 0 \\ 0 & 0 \end{bmatrix}, \tag{40}$$

$$\mathcal{H}_2^*(Y) = \frac{1}{2\sqrt{c_2}}\begin{bmatrix} R\Lambda Y R^T + (R\Lambda Y R^T)^T & 0 \\ 0 & 0 \end{bmatrix}, \tag{41}$$

$$\mathcal{H}_3^*(Y) = \frac{1}{2}\begin{bmatrix} RYR^T + (RYR^T)^T & 0 \\ 0 & 0 \end{bmatrix}. \tag{42}$$

Similarly, we define another linear operator $\mathcal{G} : \Omega_0 \to \mathbb{R}^{k\times(n-k)}$ by

$$\mathcal{G}(M,C,K) := \frac{1}{\sqrt{c_1}}(S^2)^T M_2 + \frac{1}{\sqrt{c_2}}S^T C_2 + K_2, \quad (M,C,K) \in \Omega_0. \tag{43}$$

Obviously, $\mathcal{G} : \Omega_0 \to \mathbb{R}^{k\times(n-k)}$ is surjective and its adjoint $\mathcal{G}^* : \mathbb{R}^{k\times(n-k)} \to \Omega_0$ is given by

$$\mathcal{G}^*(Z) := (\mathcal{G}_1^*(Z), \mathcal{G}_2^*(Z), \mathcal{G}_3^*(Z)), \quad Z \in \mathbb{R}^{k\times(n-k)},$$

where for each $Z \in \mathbb{R}^{k\times(n-k)}$,

$$\mathcal{G}_1^*(Z) = \frac{1}{2\sqrt{c_1}}\begin{bmatrix} 0 & S^2 Z \\ (S^2 Z)^T & 0 \end{bmatrix}, \tag{44}$$

$$\mathcal{G}_2^*(Z) = \frac{1}{2\sqrt{c_2}}\begin{bmatrix} 0 & SZ \\ (SZ)^T & 0 \end{bmatrix}, \tag{45}$$

$$\mathcal{G}_3^*(Z) = \frac{1}{2}\begin{bmatrix} 0 & Z \\ Z^T & 0 \end{bmatrix}. \tag{46}$$

Now, we define the linear operator $\mathcal{A} : \Omega_0 \to \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ by

$$\mathcal{A}(M,C,K) := (\mathcal{H}(M,C,K), \mathcal{G}(M,C,K)), \quad (M,C,K) \in \Omega_0. \tag{47}$$

Since both $\mathcal{H} : \Omega_0 \to \mathrm{Range}(\mathcal{H})$ and $\mathcal{G} : \Omega_0 \to \mathbb{R}^{k\times(n-k)}$ are surjective, $\mathcal{A} : \Omega_0 \to \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ is also surjective. The adjoint $\mathcal{A}^* : \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)} \to \Omega_0$ of $\mathcal{A}$ takes the following form

$$\mathcal{A}^*(Y,Z) = (\mathcal{A}_1^*(Y,Z), \mathcal{A}_2^*(Y,Z), \mathcal{A}_3^*(Y,Z)), \quad (Y,Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)},$$

where for each $(Y, Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$,

$$\mathcal{A}_i^*(Y, Z) = \mathcal{H}_i^*(Y) + \mathcal{G}_i^*(Z), \quad i = 1, 2, 3.$$

Let Ω_0 be equipped with the following natural inner product

$$\left\langle (M, C, K), (\widetilde{M}, \widetilde{C}, \widetilde{K}) \right\rangle := \langle M, \widetilde{M}\rangle + \langle C, \widetilde{C}\rangle + \langle K, \widetilde{K}\rangle, \quad (M, C, K), (\widetilde{M}, \widetilde{C}, \widetilde{K}) \in \Omega_0$$

and its induced norm $\|\cdot\|$. Then, problem (38), which is a special case of problem (15), becomes

$$\begin{array}{ll} \min & \frac{1}{2}\|(M, C, K) - (M_a, C_a, K_a)\|^2 \\ \text{s.t.} & \mathcal{A}(M, C, K) = 0, \\ & (M, C, K) \in \Omega. \end{array} \tag{48}$$

Define $\theta : \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)} \to \mathbb{R}$ by

$$\theta(Y, Z) := \frac{1}{2}\|\Pi_\Omega((M_a, C_a, K_a) + \mathcal{A}^*(Y, Z))\|^2 - \frac{1}{2}\|(M_a, C_a, K_a)\|^2, \tag{49}$$

where $(Y, Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$. Thus, by (22) and (19), the dual of problem (48) is

$$\begin{array}{ll} \min & \theta(Y, Z) \\ \text{s.t.} & (Y, Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}. \end{array} \tag{50}$$

Define $F : \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)} \to \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ by

$$F(Y, Z) := \nabla\theta(Y, Z) = \mathcal{A}\,\Pi_\Omega\left((M_a, C_a, K_a) + \mathcal{A}^*(Y, Z)\right), \tag{51}$$

where $(Y, Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$. We further define three mappings $\Phi_1, \Phi_2, \Phi_3 : \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)} \to \mathcal{S}^n$ by

$$\Phi_1(Y, Z) := M_a + \mathcal{A}_1^*(Y, Z),\ \Phi_2(Y, Z) := C_a + \mathcal{A}_2^*(Y, Z),\ \text{and}\ \Phi_3(Y, Z) := K_a + \mathcal{A}_3^*(Y, Z), \tag{52}$$

where $(Y, Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$. Then the mapping F defined by (51) takes the following form

$$F(Y, Z) = \nabla\theta(Y, Z) = \mathcal{A}\left(\Pi_{\mathcal{S}_+^n}(\Phi_1(Y, Z)),\, \Phi_2(Y, Z),\, \Pi_{\mathcal{S}_+^n}(\Phi_3(Y, Z))\right), \tag{53}$$

where $(Y, Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$.

In the following, we study the linear space $\mathrm{Range}(\mathcal{H})$, which varies according to the distribution of the prescribed eigenvalues. For the sake of convenience, we define a new linear operator $\mathcal{J} : \mathcal{S}^k \times \mathcal{S}^k \times \mathcal{S}^k \to \mathbb{R}^{k\times k}$ by

$$\mathcal{J}(B, D, E) := \frac{1}{\sqrt{c_1}}(\Lambda^2)^T B + \frac{1}{\sqrt{c_2}}\Lambda^T D + E, \quad (B, D, E) \in \mathcal{S}^k \times \mathcal{S}^k \times \mathcal{S}^k, \tag{54}$$

where Λ is defined by (3) and

$$k = \sum_{i=1}^{\mu} 2s_i + \sum_{i=\mu+1}^{\nu} s_i.$$

Denote

$$\mathrm{Range}(\mathcal{J}) := \{\mathcal{J}(B, D, E) \,:\, (B, D, E) \in \mathcal{S}^k \times \mathcal{S}^k \times \mathcal{S}^k\}.$$

By the definitions of Range($\mathcal{H}$) and Range($\mathcal{J}$), we can see that Range($\mathcal{H}$) = Range($\mathcal{J}$). The following result gives the exact number of $\mathrm{Dim}\,(\mathrm{Range}(\mathcal{J}))$, the dimension of Range($\mathcal{J}$).

Proposition 3.1 *Suppose that the prescribed eigen-matrix Λ is defined by* (3)*. Then we have*

$$\mathrm{Dim}\,(\mathrm{Range}(\mathcal{J})) = k^2 - \sum_{i=1}^{\mu} s_i(s_i - 1) - \frac{1}{2}\sum_{i=\mu+1}^{\nu} s_i(s_i - 1).$$

In particular, if $s_1 = \cdots = s_\mu = s_{\mu+1} = \cdots = s_\nu = 1$, $\mathrm{Dim}\,(\mathrm{Range}(\mathcal{J})) = k^2$.

Proof: For any $\widetilde{D} \in \mathcal{S}^k$ and $Y \in \mathbb{R}^{k\times k}$, denote

$$\widetilde{D} = \begin{bmatrix} \widetilde{D}_{1,1} & \cdots & \widetilde{D}_{1,\mu} & \widetilde{D}_{1,\mu+1} & \cdots & \widetilde{D}_{1,\nu} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \widetilde{D}_{1,\mu}^T & \cdots & \widetilde{D}_{\mu,\mu} & \widetilde{D}_{\mu,\mu+1} & \cdots & \widetilde{D}_{\mu,\nu} \\ \widetilde{D}_{1,\mu+1}^T & \cdots & \widetilde{D}_{\mu,\mu+1}^T & \widetilde{D}_{\mu+1,\mu+1} & \cdots & \widetilde{D}_{\mu+1,\nu} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \widetilde{D}_{1,\nu}^T & \cdots & \widetilde{D}_{\mu,\nu}^T & \widetilde{D}_{\mu+1,\nu}^T & \cdots & \widetilde{D}_{\nu,\nu} \end{bmatrix}$$

and

$$Y = \begin{bmatrix} Y_{1,1} & \cdots & Y_{1,\mu} & Y_{1,\mu+1} & \cdots & Y_{1,\nu} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Y_{\mu,1} & \cdots & Y_{\mu,\mu} & Y_{\mu,\mu+1} & \cdots & Y_{\mu,\nu} \\ Y_{\mu+1,1} & \cdots & Y_{\mu+1,\mu} & Y_{\mu+1,\mu+1} & \cdots & Y_{\mu+1,\nu} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Y_{\nu,1} & \cdots & Y_{\nu,\mu} & Y_{\nu,\mu+1} & \cdots & Y_{\nu,\nu} \end{bmatrix}, \tag{55}$$

where $\widetilde{D}_{i,j}$, $Y_{i,j} \in \mathbb{R}^{s_i\times s_j}$ and $Y_{j,i} \in \mathbb{R}^{s_j\times s_i}$ with $1 \le i \le j \le \nu$.

First, we note that for any $Y \in \mathbb{R}^{k\times k}$,

$$\begin{array}{lll} & Y \in \mathrm{Range}(\mathcal{J}) & \\ \Leftrightarrow & \frac{1}{\sqrt{c_1}}(\Lambda^2)^T B + \frac{1}{\sqrt{c_2}}\Lambda^T D + E = Y & \text{for some } (B, D, E) \in \mathcal{S}^k \times \mathcal{S}^k \times \mathcal{S}^k \\ \Leftrightarrow & E = Y - (\Lambda^2)^T(\frac{1}{\sqrt{c_1}}B) - \Lambda^T(\frac{1}{\sqrt{c_2}}D) & \text{for some } (B, D, E) \in \mathcal{S}^k \times \mathcal{S}^k \times \mathcal{S}^k \\ \Leftrightarrow & Y - (\Lambda^2)^T(\frac{1}{\sqrt{c_1}}B) - \Lambda^T(\frac{1}{\sqrt{c_2}}D) = E = E^T & \\ & \quad = Y^T - (\frac{1}{\sqrt{c_1}}B)\Lambda^2 - (\frac{1}{\sqrt{c_2}}D)\Lambda & \text{for some } (B, D, E) \in \mathcal{S}^k \times \mathcal{S}^k \times \mathcal{S}^k \\ \Leftrightarrow & (\frac{1}{\sqrt{c_2}}D)\Lambda - \Lambda^T(\frac{1}{\sqrt{c_2}}D) & \\ & \quad = Y^T - Y + (\Lambda^2)^T(\frac{1}{\sqrt{c_1}}B) - (\frac{1}{\sqrt{c_1}}B)\Lambda^2 & \text{for some } (B, D) \in \mathcal{S}^k \times \mathcal{S}^k \\ \Leftrightarrow & \widetilde{D}\Lambda - \Lambda^T\widetilde{D} = Y^T - Y & \text{for some } \widetilde{D} \in \mathcal{S}^k, \end{array} \tag{56}$$

where $\widetilde{D} = \frac{1}{\sqrt{c_2}}D + (\frac{1}{\sqrt{c_1}}B)\Lambda + \Lambda^T(\frac{1}{\sqrt{c_1}}B)$. Then, it is easy to verify that the last equation in (56) holds for some $\widetilde{D} \in \mathcal{S}^k$ if and only if

$$\widetilde{D}_{i,i}\Lambda_i - \Lambda_i^T\widetilde{D}_{i,i} = Y_{i,i}^T - Y_{i,i}, \quad i = 1, \cdots, \nu, \tag{57}$$

$$\widetilde{D}_{i,j}\Lambda_j - \Lambda_i^T\widetilde{D}_{i,j} = Y_{j,i}^T - Y_{i,j}, \quad 1 \le i < j \le \nu. \tag{58}$$

- Since $\sigma(\Lambda_i) \cap \sigma(\Lambda_j) = \emptyset$, for any $Y_{i,j} \in \mathbb{R}^{s_i \times s_j}$, there always exists a unique $\widetilde{D}_{i,j} \in \mathbb{R}^{s_i \times s_j}$ such that (58) holds.
- For $\mu + 1 \le i \le \nu$, since $\Lambda_i = \lambda_i I_{s_i}$, (57) holds if and only if

$$Y_{i,i}^T = Y_{i,i}. \tag{59}$$

- For $1 \le i \le \mu$,

$$\Lambda_i = \alpha_i I + \beta_i J, \quad J = \mathrm{diag}(\overbrace{J_2, \ldots, J_2}^{s_i}), \quad J_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

  So, (57) holds if and only if

$$\begin{array}{l} (Y_{i,i}(2q+1:2q+2, 2p+1:2p+2))^T - Y_{i,i}(2p+1:2p+2, 2q+1:2q+2) \\ \in \left\{ \begin{bmatrix} y_1 & y_2 \\ -y_2 & y_1 \end{bmatrix} : y_1, y_2 \in \mathbb{R} \right\}, \end{array} \tag{60}$$

  where $0 \le p < q \le s_i - 1$ and $Y_{i,i}(u_1 : u_2, v_1 : v_2)$ denotes the submatrix obtained by extracting rows u_1 through u_2 and columns v_1 through v_2 from a matrix $Y_{i,i}$.

Hence, we have obtained that $Y \in \mathrm{Range}(\mathcal{J})$ if and only if Y is of form (55) satisfying (59) and (60). Therefore,

$$\mathrm{Dim}\,(\mathrm{Range}(\mathcal{J})) = k^2 - \sum_{i=1}^{\mu} s_i(s_i - 1) - \frac{1}{2}\sum_{i=\mu+1}^{\nu} s_i(s_i - 1),$$

which implies that if $s_1 = \cdots = s_\mu = s_{\mu+1} = \cdots = s_\nu = 1$, $\mathrm{Dim}\,(\mathrm{Range}(\mathcal{J})) = k^2$. □

Remark 3.2 *From the proof of Proposition 3.1 on the dimension of* $\mathrm{Range}(\mathcal{J})$ *(and also* $\mathrm{Range}(\mathcal{H})$*), we actually get the general solution of* $\mathcal{J}(B, D, E) = 0$ *for any fixed* $B \in \mathcal{S}^k$*, and thus the general solution of* $\mathcal{A}(M, C, K) = 0$ *for any fixed* $M \in \mathcal{S}^n$*. This generalizes a corresponding result of [24] for* Λ *with simple eigenvalues only.*

Next, we discuss the following simplified version of the IQEP:

$$\begin{array}{ll} \min & \frac{1}{2}\|(M, C, K) - (M_a, C_a, K_a)\|^2 \\ \text{s.t.} & \mathcal{A}(M, C, K) = 0, \\ & (M, C, K) \in \Omega_0, \end{array} \tag{61}$$

which has been considered by many authors, for examples, see [13, 18]. We will use the optimal solution to problem (61) as the starting point to our Newton's method proposed in the next section. By Section 2, the unique solution (M^0, C^0, K^0) to problem (61) takes the following form

$$(M^0, C^0, K^0) = (M_a, C_a, K_a) + \mathcal{A}^*(Y^0, Z^0), \tag{62}$$

where $(Y^0, Z^0) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ is a solution of

$$\mathcal{A}\left(\Phi_1(Y,Z), \Phi_2(Y,Z), \Phi_3(Y,Z)\right) = 0, \quad (Y,Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)},$$

where $\Phi_1, \Phi_2, \Phi_3 : \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)} \to \mathcal{S}^n$ are defined by (52). Equivalently, $(Y^0, Z^0) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ is a solution of

$$\left(\mathcal{H}(\Phi_1(Y,Z), \Phi_2(Y,Z), \Phi_3(Y,Z)),\ \mathcal{G}((\Phi_1(Y,Z), \Phi_2(Y,Z), \Phi_3(Y,Z))\right) = 0, \tag{63}$$

where $(Y, Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$. The point (Y^0, Z^0) may not be a solution to the dual problem (50) as M^0 and K^0 may fail to be positive semidefinite. Nevertheless, it may be a good guess to the solution of problem (50) and will be used as an initial point in our numerical experiments.

Partition M_a, C_a, and K_a into the same structure as M, C, and K in (37),

$$M_a =: \begin{bmatrix} (M_a)_1 & (M_a)_2 \\ (M_a)_2^T & (M_a)_4 \end{bmatrix},\ C_a := \begin{bmatrix} (C_a)_1 & (C_a)_2 \\ (C_a)_2^T & (C_a)_4 \end{bmatrix},\ K_a := \begin{bmatrix} (K_a)_1 & (K_a)_2 \\ (K_a)_2^T & (K_a)_4 \end{bmatrix},$$

where $(M_a)_1, (C_a)_1, (K_a)_1 \in \mathcal{S}^k$, $(M_a)_2$, $(C_a)_2$, $(K_a)_2 \in \mathbb{R}^{k\times(n-k)}$, and $(M_a)_4, (C_a)_4, (K_a)_4 \in \mathcal{S}^{(n-k)}$. Then we obtain for any $(Y, Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ that

$$\begin{aligned}\mathcal{H}(\Phi_1(Y,Z), \Phi_2(Y,Z), \Phi_3(Y,Z)) &= \mathcal{H}(M_a, C_a, K_a) + U_1 Y R^T R \\ &+\frac{1}{2}\left(\frac{1}{c_1}(\Lambda^2)^T R^T R Y^T (\Lambda^2)^T R^T R + \frac{1}{c_2}\Lambda^T R^T R Y^T \Lambda^T R^T R + R^T R Y^T R^T R\right)\end{aligned}$$

and

$$\mathcal{G}(\Phi_1(Y,Z), \Phi_2(Y,Z), \Phi_3(Y,Z)) = \mathcal{G}(M_a, C_a, K_a) + U_2 Z,$$

where

$$U_1 := \frac{1}{2}\left(\frac{1}{c_1}(\Lambda^2)^T R^T R \Lambda^2 + \frac{1}{c_2}\Lambda^T R^T R \Lambda + R^T R\right) \tag{64}$$

and

$$U_2 := \frac{1}{2}\left(\frac{1}{c_1}(S^2)^T S^2 + \frac{1}{c_2} S^T S + I\right). \tag{65}$$

We may use the conjugate gradient (CG) method[1] [19, Algorithm 10.2.1] to solve the first equation of (63), i.e.,

$$\begin{aligned}&-\mathcal{H}(M_a, C_a, K_a) = \\ &U_1 Y R^T R + \frac{1}{2}\left(\frac{1}{c_1}(\Lambda^2)^T R^T R Y^T (\Lambda^2)^T R^T R + \frac{1}{c_2}\Lambda^T R^T R Y^T \Lambda^T R^T R + R^T R Y^T R^T R\right)\end{aligned} \tag{66}$$

[1]Strictly speaking, we should apply the CG method to the equivalent equation form $Ax = b$ of (66), where $A \in \mathcal{S}^m$, $b \in \mathbb{R}^m$, and m is the dimension of $\mathrm{Range}(\mathcal{H})$. In the implementation of CG methods there is no need to form A and b explicitly. One may also consider the preconditioned CG method to solve (66). However, cautions should be taken as the preconditioned CG method may not guarantee that the generated matrices Y stay in $\mathrm{Range}(\mathcal{H})$.

to get $Y^0 \in \text{Range}(\mathcal{H})$. The required total computational cost of computing Y^0 is $O(k^5)$ flops (the CG method needs at most k^2 steps with each step costing $O(k^3)$ flops). The solution to the second equation of (63), i.e,

$$\mathcal{G}(M_a, C_a, K_a) + U_2 Z = 0,$$

is given by

$$Z^0 = -U_2^{-1}\mathcal{G}(M_a, C_a, K_a).$$

We need $O(nk^2)$ flops to compute Z^0. The optimal solution to problem (61) can then be obtained via (62).

Remark 3.3 *The above arguments imply that if we drop the positive semidefiniteness requirement on M and K, we can obtain the solution to problem (61) with a total cost of $O(nk^2 + k^5)$. This amount of cost is small as in practice $k \ll n$. It is noted that in our approach the presence of the weighting parameters c_1 and c_2 makes no difficulty at all.*

4 Algorithm and Convergence Analysis

In this section, we shall first present a globalized version of Clarke's Jacobian based Newton method (30) for solving the dual problem (50) with $\theta : \text{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)} \to \mathbb{R}$ being defined by (49). Let $F : \text{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)} \to \text{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ be defined by (51), i.e.,

$$F(Y, Z) = \nabla\theta(Y, Z) = \mathcal{A}\,\Pi_\Omega\left((M_a, C_a, K_a) + \mathcal{A}^*(Y, Z)\right), \quad (Y, Z) \in \text{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)},$$

where $\mathcal{A} : \Omega_0 \to \text{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ is defined by (47) and $\mathcal{A}^*$ is the adjoint of $\mathcal{A}$.

Algorithm 4.1 (Newton's Method)

Step 0. *Given $(Y^0, Z^0) \in \text{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$, $\eta \in (0, 1)$, $\rho, \delta \in (0, 1/2)$. $j := 0$.*

Step 1. *Select an element $W_j \in \partial\Pi_\Omega\left((M_a, C_a, K_a) + \mathcal{A}^*(Y^j, Z^j)\right)$ and let $V_j := \mathcal{A}W_j\mathcal{A}^*$. Apply the conjugate gradient method [19, Algorithm 10.2.1] to find an approximate solution $(\Delta Y^j, \Delta Z^j) \in \text{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ to*

$$F(Y^j, Z^j) + V_j(\Delta Y, \Delta Z) = 0 \tag{67}$$

such that

$$\|F(Y^j, Z^j) + V_j(\Delta Y^j, \Delta Z^j)\| \le \eta_j \|F(Y^j, Z^j)\| \tag{68}$$

and

$$\left\langle F(Y^j, Z^j), (\Delta Y^j, \Delta Z^j)\right\rangle \le -\eta_j \left\langle (\Delta Y^j, \Delta Z^j), (\Delta Y^j, \Delta Z^j)\right\rangle, \tag{69}$$

where $\eta_j := \min\{\eta, \|F(Y^j, Z^j)\|\}$. If (68) and (69) are not achievable, let

$$(\Delta Y^j, \Delta Z^j) := -F(Y^j, Z^j) = -\mathcal{A}\,\Pi_\Omega\left((M_a, C_a, K_a) + \mathcal{A}^*(Y^j, Z^j)\right).$$

Step 2. *Let m_j be the smallest nonnegative integer m such that*

$$\theta\left((Y^j, Z^j) + \rho^m(\Delta Y^j, \Delta Z^j)\right) - \theta(Y^j, Z^j) \leq \delta\rho^m \left\langle F(Y^j, Z^j), (\Delta Y^j, \Delta Z^j)\right\rangle.$$

Set

$$(Y^{j+1}, Z^{j+1}) := (Y^j, Z^j) + \rho^{m_j}(\Delta Y^j, \Delta Z^j).$$

Step 3. *Replace j by $j+1$ and go to Step 1.*

Below we make several comments on the above algorithm before we present our convergence analysis.

- *The Starting point.* One may choose an arbitrary starting point (Y^0, Z^0) as long as $(Y^0, Z^0) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$. In numerical computation, we recommend choosing the solution to equation (63) as (Y^0, Z^0).

- *The generalized Jacobian.* In Step 1 of Algorithm 4.1, we need to select an element

$$W_j \in \partial\Pi_\Omega\left((M_a, C_a, K_a) + \mathcal{A}^*(Y^j, Z^j)\right)$$

  at the j-th iteration. This can be computed according to (36) once the spectral decompositions of $M_a + \mathcal{A}_1^*(Y^j, Z^j)$ and $K_a + \mathcal{A}_3^*(Y^j, Z^j)$ are available.

- *The conjugate gradient method.* In Step 1 of Algorithm 4.1, if we apply the CG method [19, Algorithm 10.2.1] to (67), we can always keep $(\Delta Y^j, \Delta Z^j) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ when the initial guess is in $\mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$. On may also consider the preconditioned CG method to solve (66) as long as the preconditioned CG method can guarantee that $(\Delta Y^j, \Delta Z^j) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$.

The next theorem is our global convergence result.

Theorem 4.2 *For any $(Y^0, Z^0) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$, Algorithm 4.1 generates an infinite sequence $\{(Y^j, Z^j)\}$ with the properties that for each $j \geq 0$, $(Y^j, Z^j) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$, $\{(Y^j, Z^j)\}$ is bounded, and any accumulation point of $\{(Y^j, Z^j)\}$ is a solution to problem (50).*

Proof: Since, by Step 1 of Algorithm 4.1, for any $j \geq 0$, $(\Delta Y^j, \Delta Z^j)$ is always a descent direction of $\theta(\cdot)$ at (Y^j, Z^j) and $(\Delta Y^j, \Delta Z^j) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$, Algorithm 4.1 is well defined. Then, an infinite sequence $\{(Y^j, Z^j)\}$ is generated and for each $j \geq 0$, $(Y^j, Z^j) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$.

From Theorem 1.1 we know that problem (38) has a strictly feasible solution. This, together with the surjectivity of $\mathcal{A} : \Omega_0 \to \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$, implies that the generalized Slater condition (13) holds for problem (38). Thus, by Section 2, the level set

$$\{(Y, Z) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)} \,:\, \theta(Y, Z) \leq \theta(Y^0, Z^0)\}$$

is bounded. Therefore, the sequence $\{(Y^j, Z^j)\}$ is bounded. Furthermore, by employing standard convergence analysis of optimization methods (cf. [14, Theorem 6.3.3]), we can conclude that

$$\lim_{j\to\infty} \nabla\theta(Y^j, Z^j) = 0\,,$$

which, together with the convexity of $\theta(\cdot)$ and the boundedness of $\{(Y^j, Z^j)\}$, implies that any accumulation point of $\{(Y^j, Z^j)\}$ is a solution to (50). $\square$

Theorem 4.2 shows that our algorithm converges globally. For discussions on the rate of convergence, we need the constraint nondegenerate condition (29). In our case, this condition can be written as

$$\mathcal{A}\left(\lin\left(\mathcal{T}_{\mathcal{S}^n_+}(\overline{M})\right), \mathcal{S}^n, \lin\left(\mathcal{T}_{\mathcal{S}^n_+}(\overline{K})\right)\right) = \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}, \tag{70}$$

where $(\overline{M}, \overline{C}, \overline{K}) \in \Omega_0$ is a feasible solution to problem (38). Certainly, condition (70) may fail to hold for some feasible points of problem (38). For example, if $\overline{M} = \overline{C} = \overline{K} = 0 \in \mathcal{S}^n$, then by (34), the constraint nondegenerate condition (70) is equivalent to

$$\mathcal{A}\left(0, \mathcal{S}^n, 0\right) = \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)},$$

i.e.,

$$\mathcal{H}\left(0, \mathcal{S}^n, 0\right) = \mathrm{Range}(\mathcal{H}) \quad \text{and} \quad \mathcal{G}\left(0, \mathcal{S}^n, 0\right) = \mathbb{R}^{k\times(n-k)}.$$

The relation $\mathcal{G}\left(0, \mathcal{S}^n, 0\right) = \mathbb{R}^{k\times(n-k)}$ always holds if Λ is nonsingular. The relation $\mathcal{H}\left(0, \mathcal{S}^n, 0\right) = \mathrm{Range}(\mathcal{H})$ is not true if Λ has only simple eigenvalues as in this case, by Proposition 3.1, $\mathrm{Range}(\mathcal{H}) = \mathbb{R}^{k\times k}$ while the dimension of $\mathcal{H}\left(0, \mathcal{S}^n, 0\right)$ is at most $\frac{1}{2}k \times (k+1)$.

On the other hand, condition (70) holds for many other feasible solutions. For example, if $(\overline{M}, \overline{C}, \overline{K}) \in \Omega_0$ is a feasible solution to problem (38) satisfying $\overline{K} \succ 0$, then condition (70) holds at $(\overline{M}, \overline{C}, \overline{K})$ automatically. To see this, we only need to notice from the proof of Proposition 3.1 that

$$\mathcal{H}\left(\lin\left(\mathcal{T}_{\mathcal{S}^n_+}(\overline{M})\right), \mathcal{S}^n, \mathcal{S}^n\right) = \mathrm{Range}(\mathcal{H}).$$

and

$$\mathcal{G}\left(\lin\left(\mathcal{T}_{\mathcal{S}^n_+}(\overline{M})\right), \mathcal{S}^n, \mathcal{S}^n\right) = \mathbb{R}^{k\times(n-k)}.$$

Another interesting special case to guarantee condition (70) is $\Lambda = \tau I$ for some $\tau \neq 0$. This can be seen easily because in this case, for any feasible solution $(\overline{M}, \overline{C}, \overline{K})$ to problem (38), we have

$$\mathcal{H}\left(\lin\left(\mathcal{T}_{\mathcal{S}^n_+}(\overline{M})\right), \mathcal{S}^n, \lin\left(\mathcal{T}_{\mathcal{S}^n_+}(\overline{K})\right)\right) = \mathcal{S}^k = \mathrm{Range}(\mathcal{H})$$

and

$$\mathcal{G}\left(\lin\left(\mathcal{T}_{\mathcal{S}^n_+}(\overline{M})\right), \mathcal{S}^n, \lin\left(\mathcal{T}_{\mathcal{S}^n_+}(\overline{K})\right)\right) = \mathbb{R}^{k\times(n-k)} = \mathrm{Range}(\mathcal{G}).$$

Next, we state the quadratic convergence of Algorithm 4.1.

Theorem 4.3 *Let $(\overline{Y}, \overline{Z})$ be an accumulation point of the infinite sequence $\{(Y^j, Z^j)\}$ generated by Algorithm 4.1 for solving the dual problem (50). Let*

$$(\overline{M}, \overline{C}, \overline{K}) := \Pi_\Omega\left((M_a, C_a, K_a) + \mathcal{A}^*(\overline{Y}, \overline{Z})\right).$$

Assume that the constraint nondegenerate condition (70) holds at $(\overline{M}, \overline{C}, \overline{K})$. Then the whole sequence $\{(Y^j, Z^j)\}$ converges to $(\bar{Y}, \bar{Z})$ quadratically.

Proof: By Theorem 4.2, $(\overline{Y},\overline{Z})$ is a solution of the dual problem (50) such that

$$F(\overline{Y},\overline{Z}) = \nabla\theta(\overline{Y},\overline{Z}) = 0.$$

From the the constraint nondegenerate condition (70) and Propositions 2.7 and 2.9, we know that any element $V := \mathcal{A}W\mathcal{A}^*$ with $W \in \partial\Pi_\Omega(\overline{M},\overline{C},\overline{K}))$ is positive definite. Then, by Proposition 2.3, any element in $\partial F(\overline{Y},\overline{Z})$ is also positive definite. This, together with the convexity of θ, implies that $(\overline{Y},\overline{Z})$ is the unique solution of (50). Consequently, by Theorem 4.2, the whole sequence $\{(Y^j,Z^j)\}$ converges to $(\overline{Y},\overline{Z})$. Thus, for all j sufficiently large, V_j is positive definite and $\{\|V_j^{-1}\|\}$ is uniformly bounded. Hence, for all j sufficiently large, the CG method can find $(\Delta Y^j,\Delta Z^j) \in \mathrm{Range}(\mathcal{H}) \times \mathbb{R}^{k\times(n-k)}$ such that both (68) and (69) are satisfied. Then, by Propositions 2.8 and 2.9, for all j sufficiently large,

$$\begin{array}{rcl} \|(Y^j,Z^j)+(\Delta Y^j,\Delta Z^j)-(\overline{Y},\overline{Z})\| & \le & O(\|(Y^j,Z^j)-(\overline{Y},\overline{Z})\|^2)+\eta_j\|V_j^{-1}\|\|F(Y^j,Z^j)\| \\ & = & O(\|(Y^j,Z^j)-(\overline{Y},\overline{Z})\|^2)+O(\|F(Y^j,Z^j)\|^2) \\ & = & O(\|(Y^j,Z^j)-(\overline{Y},\overline{Z})\|^2). \end{array} \tag{71}$$

Therefore, since F is strongly semismooth at $(\overline{Y},\overline{Z})$ and every element in $\partial F(\overline{Y},\overline{Z})$ is nonsingular, we know from [16, Theorem 3.3] that for all j sufficiently large,

$$\theta\left((Y^j,Z^j)+(\Delta Y^j,\Delta Z^j)\right)-\theta(Y^j,Z^j) \le \delta\left\langle F(Y^j,Z^j),(\Delta Y^j,\Delta Z^j)\right\rangle,$$

which implies that for all j sufficiently large,

$$(Y^{j+1},Z^{j+1}) = (Y^j,Z^j)+(\Delta Y^j,\Delta Z^j).$$

This, together with (71), completes the proof. □

Theorems 4.2 and 4.3 give global and local convergence analysis of Algorithm 4.1 for solving the IQEP under Assumption A.1. In fact, our algorithm can also be used to solve the IQEP of general matrix weights even Assumption A.1 fails to hold. Without giving details, below we briefly outline the way of doing so.

Let $\Phi_M,\Phi_C,\Phi_K \in \mathcal{S}^n$ be three positive definite matrices. The IQEP of matrix weights can be define as

$$\begin{array}{ll} \min & \frac{1}{2}\|\Phi_M^{-1}(M-M_a)\Phi_M^{-1}\|^2+\frac{1}{2}\|\Phi_C^{-1}(C-C_a)\Phi_C^{-1}\|^2+\frac{1}{2}\|\Phi_K^{-1}(K-K_a)\Phi_K^{-1}\|^2 \\ \text{s.t.} & M\begin{bmatrix} R \\ 0 \end{bmatrix}\Lambda^2+C\begin{bmatrix} R \\ 0 \end{bmatrix}\Lambda+K\begin{bmatrix} R \\ 0 \end{bmatrix}=0, \\ & (M,C,K)\in\Omega. \end{array} \tag{72}$$

By renaming $M := \Phi_M^{-1}M\Phi_M^{-1}$, $C := \Phi_C^{-1}C\Phi_C^{-1}$, $K := \Phi_K^{-1}K\Phi_K^{-1}$, $M_a := \phi_M^{-1}M_a\Phi_M^{-1}$, $C_a := \Phi_C^{-1}C_a\Phi_C^{-1}$, and $K_a := \Phi_K^{-1}K_a\Phi_K^{-1}$, we can write problem (72) equivalently as

$$\begin{array}{ll} \min & \frac{1}{2}\|M-M_a\|^2+\frac{1}{2}\|C-C_a\|^2+\frac{1}{2}\|K-K_a\|^2 \\ \text{s.t.} & (\Phi_M M\Phi_M)\begin{bmatrix} R \\ 0 \end{bmatrix}\Lambda^2+(\Phi_C C\Phi_C)\begin{bmatrix} R \\ 0 \end{bmatrix}\Lambda+(\Phi_K K\Phi_K)\begin{bmatrix} R \\ 0 \end{bmatrix}=0, \\ & (M,C,K)\in\Omega. \end{array} \tag{73}$$

Then, from the proof of Theorem 1.1, we can see that problem (73) admits a strictly feasible solution if and only if Λ is nonsingular.

Next, we shall show how problem (72) can be reduced to a similar problem which has a strictly feasible solution if Λ is singular. Without loss of generality, we assume that

$$\Lambda = \mathrm{diag}\{O_t, \Gamma\}, \tag{74}$$

where $O_t = \mathrm{diag}\{\overbrace{0, \ldots, 0}^{t}\}$ and Γ is nonsingular and has the same structure as Λ defined in (3). Partition M, C, K, and $\begin{bmatrix} R \\ 0 \end{bmatrix}$ by

$$M = \begin{bmatrix} M_1 & M_2 \\ M_2^T & M_4 \end{bmatrix}, C = \begin{bmatrix} C_1 & C_2 \\ C_2^T & C_4 \end{bmatrix}, K = \begin{bmatrix} K_1 & K_2 \\ K_2^T & K_4 \end{bmatrix}, \text{ and } \begin{bmatrix} R \\ 0 \end{bmatrix} = \begin{bmatrix} R_1 & R_2 \\ 0 & R_3 \end{bmatrix}, \tag{75}$$

where $M_1, C_1, K_1 \in \mathcal{S}^t$, $M_2, K_2, C_2, \in \mathbb{R}^{t\times(n-t)}$, $M_4, C_4, K_4 \in \mathcal{S}^{(n-t)}$, and $R_1 \in \mathbb{R}^{t\times t}$, $R_2 \in \mathbb{R}^{t\times(k-t)}$, $R_3 \in \mathbb{R}^{(n-t)\times(k-t)}$. The matrix R_3 can be further written as $R_3 = \begin{bmatrix} R_4 \\ 0 \end{bmatrix}$, where $R_4 \in \mathbb{R}^{(k-t)\times(k-t)}$ is an invertible upper triangular matrix. Then the first equation in problem (72) takes the following form

$$\begin{cases} K_1R_1 = 0, \quad K_2^TR_1 = 0, \\ (M_1R_2\Gamma^2 + C_1R_2\Gamma) + (M_2R_3\Gamma^2 + C_2R_3\Gamma) + (K_1R_2 + K_2R_3) = 0, \\ (M_2^TR_2\Gamma^2 + C_2^TR_2\Gamma) + (M_4R_3\Gamma^2 + C_4R_3\Gamma + K_4R_3) + K_2^TR_2 = 0. \end{cases} \tag{76}$$

Since R_1 is invertible, we have

$$K_1 = 0 \quad \text{and} \quad K_2 = 0.$$

Let

$$\widehat{\mathcal{A}}(M, C, K_4) := \begin{pmatrix} (M_1R_2\Gamma^2 + C_1R_2\Gamma) + (M_2R_3\Gamma^2 + C_2R_3\Gamma) \\ (M_2^TR_2\Gamma^2 + C_2^TR_2\Gamma) + (M_4R_3\Gamma^2 + C_4R_3\Gamma + K_4R_3) \end{pmatrix}.$$

Then the last two equations in (76) reduce to $\widehat{\mathcal{A}}(M, C, K_4) = 0$. Partition K_a by

$$K_a = \begin{bmatrix} (K_a)_1 & (K_a)_2 \\ (K_a)_2^T & (K_a)_4 \end{bmatrix}$$

with $(K_a)_1 \in \mathcal{S}^t$, $(K_a)_2 \in \mathbb{R}^{t\times(n-t)}$, and $(K_a)_4 \in \mathcal{S}^{(n-t)}$. Let

$$\Psi := \Phi_K^{-1} = \begin{bmatrix} L_1 & L_2 \\ L_2^T & L_4 \end{bmatrix} \quad \text{and} \quad \Psi K_a \Psi := \begin{bmatrix} (H_a)_1 & (H_a)_2 \\ (H_a)_2^T & (H_a)_4 \end{bmatrix}$$

where $L_1, (H_a)_1 \in \mathcal{S}^t$, $L_2, (H_a)_2 \in \mathbb{R}^{t\times(n-t)}$, and $L_4, (H_a)_4 \in \mathcal{S}^{(n-t)}$. By using the fact that

$$\begin{aligned} \|\Psi(K - K_a)\Psi\|^2 &= \left\| \begin{bmatrix} L_2K_4L_2^T - (H_a)_1 & L_2K_4L_4 - (H_a)_2 \\ L_4K_4L_2^T - (H_a)_2^T & L_4K_4L_4 - (H_a)_4 \end{bmatrix} \right\|^2 \\ &= \left\|\Psi_K^{-1}(K_4 - \Upsilon_a)\Psi_K^{-1}\right\|^2 + \kappa, \end{aligned}$$

where

$$\begin{aligned}\Psi_K &= (L_2^T L_2 + L_4^2)^{-\frac{1}{2}},\\ \Upsilon_a &= \Psi_K^2\left[(L_2^T(H_a)_1 L_2 + 2L_2^T(H_a)_2 L_4 + L_4(H_a)_4 L_4\right]\Psi_K^2,\\ \kappa &= -\|\Psi_K^{-1}\Upsilon_a\Psi_K^{-1}\|^2 + \|\Psi K_a \Psi\|^2,\end{aligned}$$

we can see that problem (72) is equivalent to

$$\begin{array}{ll}\min & \frac{1}{2}\|\Phi_M^{-1}(M-M_a)\Phi_M^{-1}\|^2 + \frac{1}{2}\|\Phi_C^{-1}(C-C_a)\Phi_C^{-1}\|^2 + \frac{1}{2}\|\Psi_K^{-1}(K_4-\Upsilon_a)\Psi_K^{-1}\|^2\\ \text{s.t.} & \widehat{\mathcal{A}}(M,C,K_4) = 0,\\ & M \succeq 0, \quad C^T = C, \quad K_4 \succeq 0.\end{array} \tag{77}$$

By renaming $M := \Phi_M^{-1}M\Phi_M^{-1}$, $C := \Phi_C^{-1}C\Phi_C^{-1}$, $K_4 := \Psi_K^{-1}K_4\Psi_K^{-1}$, $M_a := \Phi_M^{-1}M_a\Phi_M^{-1}$, $C_a := \Phi_C^{-1}C_a\Phi_C^{-1}$, and $\Upsilon_a := \Psi_K^{-1}\Upsilon_a\Psi_K^{-1}$, we see that problem (77) takes the following form

$$\begin{array}{ll}\min & \frac{1}{2}\|M-M_a\|^2 + \frac{1}{2}\|C-C_a\|^2 + \frac{1}{2}\|K_4-\Upsilon_a\|^2\\ \text{s.t.} & \widehat{\mathcal{A}}\,(\Phi_M M\Phi_M, \Phi_C C\Phi_C, \Psi_K K_4 \Psi_K) = 0\\ & M \succeq 0, \quad C^T = C, \quad K_4 \succeq 0.\end{array} \tag{78}$$

Let $\widehat{M}$, $\widehat{C}$, and $\widehat{K}_4$ be defined by

$$\widehat{M} := \begin{bmatrix}\widehat{M}_1 & \widehat{M}_2\\ \widehat{M}_2^T & \widehat{M}_4\end{bmatrix},\ \widehat{C} := \begin{bmatrix}\widehat{C}_1 & \widehat{C}_2\\ \widehat{C}_2^T & \widehat{C}_4\end{bmatrix}, \quad \text{and} \quad \widehat{K}_4 := \begin{bmatrix}R_4^{-T}\Gamma^T\Gamma R_4^{-1} & 0\\ 0 & I\end{bmatrix}, \tag{79}$$

where

$$\widehat{M}_1 = R_1^{-T}R_1^{-1}, \quad \widehat{M}_2 = -\begin{bmatrix}R_1^{-T}R_1^{-1}R_2R_4^{-1} & 0\end{bmatrix}, \quad \widehat{C}_1 = 0, \quad \widehat{C}_2 = 0$$

and

$$\widehat{M}_4 = \begin{bmatrix}R_4^{-T}R_2^TR_1^{-T}R_1^{-1}R_2R_4^{-1} + R_4^{-T}R_4^{-1} & 0\\ 0 & I\end{bmatrix}, \quad \widehat{C}_4 = -\begin{bmatrix}R_4^{-T}(\Gamma+\Gamma^T)R_4^{-1} & 0\\ 0 & 0\end{bmatrix}.$$

Clearly,

$$\left(\Phi_M^{-1}\widehat{M}\Phi_M^{-1}, \Phi_C^{-1}\widehat{C}\Phi_C^{-1}, \Psi_K^{-1}\widehat{K}_4\Psi_K^{-1}\right)$$

is a strictly feasible solution to problem (78). Therefore, we can apply Algorithm 4.1 to the dual of problem (78).

5 Numerical Experiments

In this section, we report our numerical experiments of Algorithm 4.1 for solving the IQEP (38) carried out in `MATLAB 7.0.1` running on a PC Intel Pentium IV of 2.40 GHz CPU. In both of the following two numerical examples, we randomly generate $\Lambda \in \mathbb{R}^{k\times k}$ by using the built-in function `randn` in `MATLAB 7.0.1`. The total number of complex-valued eigenvalues is chosen to be around $k/2$. The matrix $R \in \mathbb{R}^{k\times k}$ is obtained by applying the QR factorization to a random generated $n \times k$ matrix X by using `randn`.

Example 5.1 *Let $\overline{M}$, $\overline{C}$ and $\overline{K}$ be given by (8). Set*

$$M_a := \overline{M} + \tau R_M, \quad C_a := \overline{C} + \tau R_C, \quad K_a := \overline{K} + tR_K,$$

where R_M, R_C, and R_K are $n \times n$ are symmetric matrices with random entries uniformly distributed between -1.0 and 1.0 and $\tau \in \mathbb{R}$ is a perturbed parameter. We report our numerical results for (a) $k = 30$, $n = 100, 200, 500, 1000, 1500, 2000$, and $\tau = 0.1, 1.0$ and (b) $k \approx n/3$, $n = 100, 200, 300, 400, 450$, and $\tau = 0.1, 1.0$.

Example 5.2 *The matrices G_M and G_K are random $n \times n$ correlation matrices generated by `MATLAB 7.0.1's gallery ('randcorr',n)`, and the matrix G_C is a random $n \times n$ symmetric matrix with entries $(G_C)_{ij} \in [-1, 1]$ and $(G_C)_{ii} = 1.0$ for $i, j = 1, 2, \ldots, n$. Then M_a, C_a and K_a are, respectively, obtained via perturbing $\sqrt{c_1}G_M$, $\sqrt{c_2}G_C$, and G_K by a random $n \times n$ symmetric matrix with entries in $[-\tau, \tau]$, where $\tau = 0.1, 1.0$. We report our numerical results for (a) $k = 30$, $n = 100, 200, 500, 1000, 1500, 2000$, and $\tau = 0.1, 1.0$ and (b) $k \approx n/3$, $n = 100, 200, 300, 400, 450$, and $\tau = 0.1, 1.0$.*

In our numerical experiments, the initial point (Y^0, Z^0) is chosen to be the solution to equation (63) and the stopping criterion is

$$\texttt{Tol.} := \frac{\|\nabla\theta(Y_k, Z_k)\|}{\max\left\{1, \left\|\left(\frac{1}{\sqrt{c_1}}M_a, \frac{1}{\sqrt{c_2}}C_a, K_a\right)\right\|\right\}} \leq 10^{-7}\,.$$

We set other parameters used in our algorithm as $\eta = 10^{-6}$, $\rho = 0.5$, and $\delta = 10^{-4}$. Our numerical results are given in Tables 1–4, where `It.`, `Func.`, `Res0.` and `Res*.` stand for the number of iterations, the number of function evaluations, and the residuals $\|\nabla\theta(\cdot)\|$ at the starting point (Y^0, Z^0) and the final iterate of our algorithm, respectively. In Algorithm 4.1, the major cost at the j-th iteration is solving the linear system (67), i.e.,

$$\begin{pmatrix} \mathcal{H}\left(\Pi_\Omega((M_a, C_a, K_a) + \mathcal{A}^*(Y^j, Z^j)) + W_j\mathcal{A}^*(\Delta Y^j, \Delta Z^j)\right) \\ \mathcal{G}\left(\Pi_\Omega((M_a, C_a, K_a) + \mathcal{A}^*(Y^j, Z^j)) + W_j\mathcal{A}^*(\Delta Y^j, \Delta Z^j)\right) \end{pmatrix} = 0,$$

where $W_j \in \partial\Pi_\Omega\left((M_a, C_a, K_a) + \mathcal{A}^*(Y^j, Z^j)\right)$. For better numerical performance, instead of solving the above linear system directly we apply the CG method to the following linear system

$$\begin{pmatrix} U_1^{-1}\mathcal{H}\left(\Pi_\Omega((M_a, C_a, K_a) + \mathcal{A}^*(Y^j, Z^j)) + W_j\mathcal{A}^*(\Delta Y^j, \Delta Z^j)\right)(R^TR)^{-1} \\ U_2^{-1}\mathcal{G}\left(\Pi_\Omega((M_a, C_a, K_a) + \mathcal{A}^*(Y^j, Z^j)) + W_j\mathcal{A}^*(\Delta Y^j, \Delta Z^j)\right) \end{pmatrix} = 0,$$

where U_1 and U_2 are given by (64) and (65), respectively. That is, we apply a preconditioned CG method to the linear system (67).

Based on our numerical experiments, we make the following observations.

- Our algorithm converges to the required accuracy at most dozens of iterations for all but one case ($n = 2,000$, $k = 30$, $c_1 = 10$, $c_2 = 0.1$, $\tau = 1$) in Table 3. Quadratic convergence was observed for all convergent cases. This confirms our theoretical result on quadratic convergence.

$k = 30, \quad c_1 = c_2 = 1.0$							
τ	n	cputime	It.	Func.	Res0.	Res*.	Tol.
0.1	100	01 m 26 s	18	24	1.2×10^{2}	5.4×10^{-10}	3.9×10^{-11}
	200	04 m 39 s	14	15	3.2×10^{2}	9.0×10^{-10}	3.9×10^{-11}
	500	21 m 16 s	11	12	1.8×10^{3}	6.1×10^{-9}	1.3×10^{-10}
	1,000	44 m 13 s	9	10	4.2×10^{3}	9.3×10^{-8}	1.1×10^{-9}
	1,500	08 h 49 m 11 s	7	8	9.0×10^{3}	1.9×10^{-6}	1.6×10^{-8}
	2,000	05 h 24 m 37 s	9	10	1.4×10^{4}	5.0×10^{-6}	3.3×10^{-8}
1.0	100	40.6 s	10	11	1.2×10^{3}	1.9×10^{-9}	2.7×10^{-11}
	200	01 m 37 s	9	11	3.7×10^{3}	1.8×10^{-6}	1.3×10^{-8}
	500	09 m 03 s	10	12	1.4×10^{4}	7.0×10^{-7}	2.0×10^{-9}
	1,000	01 h 09 m 50 s	10	11	4.5×10^{4}	1.7×10^{-5}	2.4×10^{-8}
	1,500	08 h 12 m 53 s	14	18	6.7×10^{4}	3.9×10^{-5}	3.6×10^{-8}
	2,000	08 h 54 m 57 s	10	12	1.0×10^{5}	8.0×10^{-5}	5.7×10^{-8}
$k = 30, \quad c_1 = 10.0 \quad \text{and} \quad c_2 = 0.1$							
τ	n	cputime	It.	Func.	Res0.	Res*.	Tol.
0.1	100	35.5 s	9	10	7.7×10^{1}	5.1×10^{-9}	2.8×10^{-10}
	200	04 m 18 s	9	10	2.8×10^{2}	1.7×10^{-6}	5.1×10^{-8}
	500	20 m 14 s	7	8	1.4×10^{3}	5.4×10^{-7}	7.3×10^{-9}
	1,000	53 m 30 s	8	9	3.7×10^{3}	1.8×10^{-6}	1.3×10^{-8}
	1,500	05 h 50 m 22 s	10	11	6.0×10^{3}	7.9×10^{-6}	3.7×10^{-8}
	2,000	05 h 22 m 25 s	9	10	8.1×10^{3}	1.3×10^{-6}	4.8×10^{-9}
1.0	100	36.4 s	6	7	1.3×10^{3}	1.3×10^{-6}	9.4×10^{-9}
	200	03 m 10 s	7	8	3.7×10^{3}	5.1×10^{-8}	1.9×10^{-10}
	500	19 m 33 s	8	9	1.2×10^{4}	5.3×10^{-6}	7.8×10^{-9}
	1,000	02 h 02 m 35 s	10	11	2.8×10^{4}	8.9×10^{-7}	6.6×10^{-10}
	1,500	04 h 52 m 15 s	11	14	6.9×10^{4}	6.7×10^{-7}	3.3×10^{-10}
	2,000	12 h 30 m 15 s	11	12	8.9×10^{4}	3.2×10^{-5}	1.2×10^{-8}

Table 1: Numerical results of Example 5.1 (a)

- The largest numerical examples that we tested in this paper are: (i) $n = 2,000$ and $k = 30$ and (ii) $n = 450$ and $k = 150$. For case (i), there are roughly $6,000,000$ unknowns in the primal problem and $60,000$ unknowns in the dual problem while for case (ii), these numbers are roughly $300,000$ and $67,000$, respectively. In consideration of the scales of problems solved, our algorithm is very effective.

- We also tested the cases for $k \approx n/4$, which have better numerical performance than the cases that $k \approx n/3$. Generally speaking, the smaller the ration $\frac{k}{n}$ is, the better our algorithm performs. This is no surprise as a smaller ration $\frac{k}{n}$ implies that the constraint nondegenerate condition (70) is more likely to hold.

$k \approx n/3, \quad c_1 = c_2 = 1.0$								
τ	n	k	cputime	It.	Func.	Res0.	Res*.	Tol.
0.1	100	33	02 m 31 s	18	23	1.0×10^{2}	9.8×10^{-7}	7.2×10^{-8}
	200	66	01 h 05 m 26 s	40	69	5.3×10^{2}	1.4×10^{-8}	6.3×10^{-10}
	300	100	01 h 37 m 32 s	20	21	1.0×10^{3}	4.8×10^{-7}	1.6×10^{-8}
	400	133	10 h 23 m 23 s	15	16	2.0×10^{3}	1.6×10^{-8}	4.5×10^{-10}
	450	150	10 h 24 m 04 s	14	15	2.7×10^{3}	3.0×10^{-8}	7.4×10^{-10}
1.0	100	33	01 m 43 s	10	12	1.4×10^{3}	1.7×10^{-7}	2.3×10^{-9}
	200	66	22 m 02 s	19	35	5.5×10^{3}	1.4×10^{-5}	9.5×10^{-8}
	300	100	01 h 32 m 58 s	12	14	1.2×10^{4}	1.3×10^{-5}	5.9×10^{-8}
	400	133	02 h 46 m 54 s	14	16	2.4×10^{4}	2.4×10^{-6}	8.3×10^{-9}
	450	150	09 h 38 m 15 s	18	21	2.9×10^{4}	4.9×10^{-7}	1.5×10^{-9}
$k \approx n/3, \quad c_1 = 10.0 \quad \text{and} \quad c_2 = 0.1$								
τ	n	k	cputime	It.	Func.	Res0.	Res*.	Tol.
0.1	100	33	46.1 s	9	11	7.5×10^{1}	2.6×10^{-8}	1.4×10^{-9}
	200	66	42 m 42 s	13	15	3.6×10^{2}	1.8×10^{-6}	5.8×10^{-8}
	300	100	02 h 24 m 23 s	17	20	1.1×10^{3}	2.9×10^{-7}	6.5×10^{-9}
	400	133	04 h 38 m 42 s	10	11	1.8×10^{3}	2.4×10^{-6}	4.0×10^{-8}
	450	150	12 h 23 m 44 s	13	14	2.2×10^{3}	5.8×10^{-7}	8.8×10^{-9}
1.0	100	33	02 m 29 s	10	11	1.5×10^{3}	2.5×10^{-8}	1.8×10^{-10}
	200	66	23 m 26 s	10	11	5.0×10^{3}	5.5×10^{-8}	2.0×10^{-10}
	300	100	01 h 17 m 47 s	10	11	1.0×10^{4}	2.9×10^{-7}	7.1×10^{-10}
	400	133	06 h 24 m 49 s	17	22	1.9×10^{4}	8.4×10^{-6}	1.5×10^{-8}
	450	150	06 h 39 m 33 s	11	12	2.6×10^{4}	6.7×10^{-7}	1.1×10^{-9}

Table 2: Numerical results of Example 5.1 (b)

$k = 30, \quad c_1 = c_2 = 1.0$							
τ	n	cputime	It.	Func.	Res0.	Res*.	Tol.
0.1	100	25.1 s	9	10	6.2×10^{2}	1.2×10^{-8}	2.6×10^{-10}
	200	02 m 17 s	8	9	2.4×10^{3}	8.9×10^{-9}	1.0×10^{-10}
	500	16 m 55 s	10	53	9.9×10^{3}	6.5×10^{-8}	3.1×10^{-10}
	1,000	01 h 39 m 23 s	11	12	2.8×10^{4}	1.9×10^{-5}	4.6×10^{-8}
	1,500	04 h 48 m 57 s	11	27	5.5×10^{4}	3.9×10^{-5}	6.2×10^{-8}
	2,000	05 h 01 m 57 s	7	8	8.7×10^{4}	1.9×10^{-5}	2.2×10^{-8}
1.0	100	40.3 s	13	15	1.7×10^{3}	1.3×10^{-7}	1.5×10^{-9}
	200	01 m 30 s	11	13	4.2×10^{3}	1.4×10^{-6}	8.4×10^{-9}
	500	31 m 46 s	15	45	1.3×10^{4}	2.5×10^{-8}	6.1×10^{-11}
	1,000	03 h 04 m 28 s	12	15	4.7×10^{4}	1.8×10^{-7}	2.2×10^{-10}
	1,500	05 h 09 m 11 s	10	11	8.7×10^{4}	1.6×10^{-5}	1.3×10^{-8}
	2,000	12 h 47 m 19 s	9	10	1.2×10^{5}	2.3×10^{-5}	1.4×10^{-8}
$k = 30, \quad c_1 = 10.0$ and $c_2 = 0.1$							
τ	n	cputime	It.	Func.	Res0.	Res*.	Tol.
0.1	100	31.4 s	8	9	1.7×10^{2}	9.5×10^{-9}	2.0×10^{-10}
	200	04 m 30 s	11	12	3.0×10^{2}	5.3×10^{-6}	5.9×10^{-8}
	500	17 m 31 s	8	9	1.8×10^{3}	1.8×10^{-6}	8.2×10^{-9}
	1,000	02 h 17 m 39 s	9	10	2.8×10^{3}	1.9×10^{-6}	4.4×10^{-9}
	1,500	04 h 32 m 10 s	7	8	6.4×10^{3}	3.5×10^{-7}	5.4×10^{-10}
	2,000	Failed	7	8	1.0×10^{4}	5.6×10^{-4}	6.5×10^{-7}
1.0	100	02 m 36 s	13	14	1.2×10^{3}	7.0×10^{-7}	4.8×10^{-9}
	200	03 m 30 s	7	8	2.8×10^{3}	3.0×10^{-8}	1.0×10^{-10}
	500	45 m 44 s	11	14	1.2×10^{4}	7.0×10^{-6}	9.8×10^{-9}
	1,000	54 m 36 s	9	11	2.9×10^{4}	1.4×10^{-5}	9.6×10^{-9}
	1,500	04 h 53 m 19 s	8	9	6.1×10^{4}	9.8×10^{-5}	4.6×10^{-8}
	2,000	10 h 53 m 54 s	6	7	9.3×10^{4}	2.2×10^{-4}	7.7×10^{-8}

Table 3: Numerical results of Example 5.2 (a)

6 Conclusions

In this paper, we introduced a quadratically convergent Newton's method for solving the IQEP with the positive semidefiniteness condition on matrices M and K. Our numerical experiments showed that our method is very efficient. We are not aware of any other algorithm that is capable of achieving the numerical results reported here. Since the vast majority of our computer cputime is spent on the preconditioned conjugate gradient method for solving the linear equation (67), it would save much computing time by finding a better preconditioner for (67). We leave this as our future research topic. Another interesting topic is to study the IQEP with M, C, and K being of various special structures.

$k \approx n/3, \quad c_1 = c_2 = 1.0$								
τ	n	k	cputime	It.	Func.	Res0.	Res*.	Tol.
0.1	100	33	47.8 s	12	13	8.3×10^{2}	1.5×10^{-6}	3.3×10^{-8}
	200	66	12 m 01 s	16	22	3.3×10^{3}	1.6×10^{-8}	1.8×10^{-10}
	300	100	01 h 58 m 22 s	14	15	7.7×10^{3}	2.2×10^{-7}	1.7×10^{-9}
	400	133	05 h 39 m 37 s	17	18	1.6×10^{4}	6.2×10^{-8}	3.7×10^{-10}
	450	150	05 h 28 m 26 s	20	21	1.7×10^{4}	5.1×10^{-8}	2.7×10^{-10}
1.0	100	33	37.4 s	9	10	1.6×10^{3}	2.1×10^{-6}	2.5×10^{-8}
	200	66	01 h 24 m 43 s	45	75	5.2×10^{3}	1.8×10^{-7}	1.1×10^{-9}
	300	100	01 h 38 m 58 s	29	33	1.5×10^{4}	6.3×10^{-7}	2.6×10^{-9}
	400	133	04 h 36 m 30 s	16	18	2.5×10^{4}	2.6×10^{-6}	8.0×10^{-9}
	450	150	04 h 55 m 35 s	18	20	3.1×10^{4}	1.0×10^{-5}	2.7×10^{-8}
$k \approx n/3, \quad c_1 = 10.0 \quad \text{and} \quad c_2 = 0.1$								
τ	n	k	cputime	It.	Func.	Res0.	Res*.	Tol.
0.1	100	33	52.0 s	8	9	1.8×10^{2}	1.0×10^{-7}	2.2×10^{-9}
	200	66	11 m 01 s	9	10	4.5×10^{2}	5.6×10^{-7}	6.3×10^{-9}
	300	100	02 h 20 m 24 s	13	14	1.1×10^{3}	1.4×10^{-7}	1.1×10^{-9}
	400	133	03 h 06 m 49 s	10	11	3.0×10^{3}	1.3×10^{-7}	7.5×10^{-10}
	450	150	06 h 01 m 45 s	15	16	2.4×10^{3}	1.3×10^{-6}	6.6×10^{-9}
1.0	100	33	01 m 54 s	9	10	9.8×10^{2}	1.7×10^{-8}	1.2×10^{-10}
	200	66	22 m 27 s	8	9	4.7×10^{3}	2.2×10^{-5}	7.6×10^{-8}
	300	100	01 h 21 m 14 s	11	13	9.8×10^{3}	4.3×10^{-6}	1.0×10^{-8}
	400	133	03 h 22 m 32 s	12	14	1.9×10^{4}	2.5×10^{-5}	4.4×10^{-8}
	450	150	06 h 59 m 29 s	16	20	2.3×10^{4}	1.7×10^{-5}	2.5×10^{-8}

Table 4: Numerical results of Example 5.2 (b)

References

[1] F. Alizadeh, J.-P.A. Haeberly, and M.L. Overton, *Complementarity and nondegeneracy in semidefinite programming*, Mathematical Programming, 77 (1997), pp. 111–128.

[2] M. Baruch, *Optimization procedure to correct stiffness and flexibility matrices using vibration data*, AIAA Journal, 16 (1978), pp. 1208–1210.

[3] M. Baruch and I.Y. Bar-Itzhack, *Optimal weighted orthogonalization of measured modes*, AIAA Journal, 16 (1978), pp. 346–351.

[4] A. Berman, *Comment on "Optimal weighted orthogonalization of measured modes"*, AIAA Journal, 17 (1979), pp. 927–928.

[5] A. Berman and E.J. Nagy, *Improvement of large analytical model using test data*, AIAA Journal, 21 (1983), pp. 1168–1173.

[6] R. Bhatia, *Matrix Analysis*, Springer, New York, 1997.

[7] J.F. Bonnans, R. Cominetti, and A. Shapiro, *Sensitivity analysis of optimization problems under second order regularity constraints*, Mathematics of Operations Research, 23 (1998), pp. 803–832.

[8] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[9] S. Boyd and L. Xiao, *Least-squares covariance matrix adjustment*, Technical report, Electrical Engineering Department, Standford University, June 2004.

[10] X. Chen, H.-D. Qi, and P. Tseng, *Analysis of nonsmooth symmetric matrix valued functions with applications to semidefinite complementarity problems*, SIAM Journal on Optimization 13 (2003), pp. 960–985.

[11] M.T. Chu, Y.C. Kuo, and W.W. Lin, *On inverse quadratic eigenvalue problems with partially prescribed eigenstructure*. SIAM Journal on Matrix Analysis and Applications, 25 (2004), pp. 995–1020.

[12] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983.

[13] B.N. Datta, *Finite-element model updating, eigenstructure assignment and eigenvalue embedding techniques for vibrating systems*, Mechanical Systems and Signal Processing, 16 (2002), pp. 83–96.

[14] J.E. Dennis and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, New Jersey, 1983.

[15] W.F. Donoghue, *Monotone Matrix Functions and Analytic Continuation*, Springer, New York, 1974.

[16] F. Facchinei, *Minimization of SC^1 functions and the Maratos effect*, Operations Research Letters, 17 (1995), pp. 131–137.

[17] M.I. Friswell, D.J. Inman, and D.F. Pilkeyand, *The direct updating of damping and stiffness matrices*, AIAA Journal, 36 (1998), pp. 491–493.

[18] M.I. Friswell and J.E. Mottershead, *Finite Element Model Updating in Structural Dynamics*, Kluwer Academic Publishers, 1995.

[19] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd edition, Johns Hopkins University Press, Baltimore and London, 1996.

[20] N.J. Higham, *Computing a nearest symmetric positive semidefinite matrix*, Linear Algebra and Its Applications, 103 (1988), pp. 103–118. 1988.

[21] N.J. Higham, *Computing the nearest correlation matrix – a problem from finance*, IMA Journal of Numerical Analysis, 22 (2002), pp. 329–343.

[22] M. Kojima and S. Shindo, *Extensions of Newton and quasi-Newton methods to systems of PC^1 equations*, Journal of Operations Research Society of Japan 29 (1986), pp. 352-374.

[23] B. Kummer, *Newton's method for nondifferentiable functions*, Advances in Mathematical Optimization, 114–125, Math. Res., 45, Akademie-Verlag, Berlin, 1988.

[24] Y.C. Kuo, W.W. Lin, and S.F. Xu, *On a general solution of partially described inverse quadratic eigenvalue problems and its applications*, Preprint 2005-1-001, National Center for Theoretical Sciences, National Tsing Hua University, Taiwan, 2005

[25] Y.C. Kuo, W.W. Lin, and S.F. Xu, *A new model updating method for quadratic eigenvalue problems. II*, Preprint 2005-2-001, National Center for Theoretical Sciences, National Tsing Hua University, Taiwan, 2005

[26] M. Link, *Updating of analytical models–Review of numerical procedures and application aspects*, Proc., Structural Dynamics Forum SD2000, Los Alamos, N.M., 1999.

[27] K. Löwner, *Über monotone matrixfunctionen*, Mathematische Zeitschrift 38 (1934), pp. 177–216.

[28] N.M.M. Maia and J.M.M. Silva, *Theoretical and experimental modal analysis*, Research Studies Press Ltd., Somerset, 1997.

[29] J. Malick, *A dual approach to semidefinite least-squares problems*, SIAM Journal Matrix Analysis and Applications, 26 (2004), pp. 272–284.

[30] F. Meng, D. Sun, and G. Zhao, *Semismoothness of solutions to generalized equations and the Moreau-Yosida regularization*, Mathematical Programming, Series B, to appear.

[31] R. Mifflin, *Semismoothness and semiconvex functions in constrained optimization*, SIAM Journal on Control and Optimization, 15 (1977), pp. 959–972.

[32] J.E. Mottershead and M.I. Friswell, *Model updating in structural dynamics: A survey*, Journal of Sound and Vibration, 167 (1993), pp. 347–375.

[33] J.S. Pang, D. Sun, and J. Sun, *Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems,* Mathematics of Operations Research, 28 (2003), pp. 39–63.

[34] H.-D. Qi, *Some theoretical aspects on Newton's method for constrained best interpolation*, Continuous Optimization: Current Trends and Modern Applications, pp. 23–49, Springer, New York, 2005.

[35] H.-D. Qi and D. Sun, *Quadratic convergence and numerical experiments of Newton's method for computing the nearest correlation matrix*, Preprint, Department of Mathematics, National University of Singapore, Singapore, 2005.

[36] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Mathematical Programming, 58 (1993), pp. 353–367.

[37] S.M. Robinson, *Local structure of feasible sets in nonlinear programming, Part II: Nondegeneracy*, Mathematical Programming Study, 22 (1984), pp. 217–230.

[38] S.M. Robinson, *Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity*, Mathematical Programming Study, 30 (1987), pp. 45–66.

[39] S.M. Robinson, *Constraint nondegeneracy in variational analysis*, Mathematics of Operations Research, 28 (2003), pp. 201–232.

[40] R.T. Rockafellar, *Conjugate Duality and Optimization*, SIAM, Philadelphia, 1974.

[41] R.T. Rockafellar and R.J.B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[42] A. Shapiro, *Sensitivity analysis of generalized equations*, Journal of Mathematical Sciences, 115 (2003), pp. 2554–2565.

[43] D. Sun and J. Sun, *Semismooth matrix valued functions*, Mathematics of Operations Research, 27 (2002), pp. 150–169.

[44] F. Tisseur and K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Review, 43 (2001), pp. 235–286.

[45] P. Tseng, *Merit functions for semi-definite complementarity problems*, Mathematical Programming, 83 (1998), pp. 159–185.

[46] K.C. Toh, R.H. Tütüncü, and M.J. Todd, *Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems*, Preprint, Department of Mathematics, National University of Singapore, Singapore, 2005.

[47] F.-S. Wei, *Structural dynamic model identification using vibration test data*, 7th IMAC, Las Vegas, Nevada, 1989, pp. 562–567.

[48] F.-S. Wei, *Structural dynamic model improvement using vibration test data*, AIAA Journal, 28 (1990), pp. 175–177.

[49] F.-S. Wei, *Mass and stiffness interaction effects in analytical model modification*, AIAA Journal, 28 (1990), pp. 1686–1688.

[50] E.H. Zarantonello, *Projections on convex sets in Hilbert space and spectral theory I and II.* In E.H. Zarantonello, editor, *Contributions to Nonlinear Functional Analysis*, Academic Press (New York, 1971) pp. 237–424.