Comput Optim Appl (2015) 62:815–850
DOI 10.1007/s10589-015-9757-1

The augmented Lagrangian method based on the APG strategy for an inverse damped gyroscopic eigenvalue problem

Yue Lu[1] · **Liwei Zhang**[1]

Received: 13 March 2014 / Published online: 7 May 2015

Abstract In this paper, we propose an augmented Lagrangian method based on the accelerated proximal gradient (APG) strategy for an inverse damped gyroscopic eigenvalue problem (IDGEP), which is a special case of the classical inverse quadratic eigenvalue problem. Under mild conditions, we show that the whole sequence of iterations generated by the proposed algorithm converges to the unique solution of the IDGEP. In view of the iteration-complexity, the proposed algorithm requires at most $O(\log(\varepsilon^{-1}))$ outer iterations and at most $O(\varepsilon^{-1})$ APG calls to obtain an ε-feasible and ε-optimal solution of the IDGEP. Numerical results indicate that the proposed algorithm can solve the test problems efficiently.

Keywords Inverse damped gyroscopic eigenvalue problem · Augmented Lagrangian method · Accelerated proximal gradient method · Iteration-complexity

Mathematics Subject Classification 15A18 · 65F15 · 65F18 · 65K05 · 90C22 · 90C25

1 Introduction

Throughout this paper, the following notations will be used:

✉ Yue Lu
jinjin403@sina.com

Liwei Zhang
lwzhang@dlut.edu.cn

[1] Institute of Operations Research and Control Theory, School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, People's Republic of China

- Let $\mathscr{R}^{m\times n}$ be the space of $m \times n$ matrices. Given any $X, Y \in \mathscr{R}^{m\times n}$, the standard inner product of X and Y is defined by $\langle X, Y\rangle := \mathrm{Tr}(XY^T)$ and the Frobenius norm of X is denoted by $\|X\|_F := \sqrt{\mathrm{Tr}(XX^T)}$, where $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix.
- Let $\mathscr{S}^n$ denote the space of $n \times n$ symmetric matrices. The cone of $n \times n$ positive semi-definite matrices is written by $\mathscr{S}^n_+$. Let $\mathscr{K}^n$ denote the set of all real $n \times n$ skew-symmetric matrices, that is, $X \in \mathscr{K}^n \Leftrightarrow X^T = -X$.
- We denote $\Omega := \mathscr{S}^n_+ \times \mathscr{S}^n \times \mathscr{S}^n_+ \times \mathscr{K}^n \times \mathscr{K}^n$ and $\Omega_0 := \mathscr{S}^n \times \mathscr{S}^n \times \mathscr{S}^n \times \mathscr{K}^n \times \mathscr{K}^n$. The indicator function of Ω is denoted by $\delta_\Omega(\cdot)$, that is,

$$\delta_\Omega(X) := \begin{cases} 0 & \text{if } X \in \Omega, \\ \infty & \text{otherwise.} \end{cases}$$

  Let $\Pi_\Omega(X)$ be the metric projection of X onto Ω, which is the unique optimal solution of the following problem

$$\begin{aligned} &\min \ \|X - Y\|^2 \\ &\text{s.t.} \quad Y \in \Omega. \end{aligned}$$

- For $x := (M, C, K, G, N) \in \Omega$, the norm of x is defined as follows:

$$\|x\| := \sqrt{\|M\|_F^2 + \|C\|_F^2 + \|K\|_F^2 + \|G\|_F^2 + \|N\|_F^2}.$$

For three given $n \times n$ real matrices M, C, K, the classical *quadratic eigenvalue problem* (QEP) is to find scalars λ and nonzero vectors x satisfying

$$Q(\lambda)x = 0, \tag{1.1}$$

where (λ, x) is called an *eigenpair* that consists of the eigenvalue and the eigenvector with respect to the *pencil* $Q(\lambda) := \lambda^2 M + \lambda C + K$. In the past decade, QEPs have gained remarkable attention in various engineering applications such as applied mechanics, electrical oscillation, fluid mechanics, signal processing, vibro-acoustics and finite element models of some PDEs. Usually, the matrices M, C and K in the pencil $Q(\lambda)$ are required to be symmetric. In particular, the finite element model (FEM) [1] is of the form

$$M\ddot{x}(t) + C\dot{x}(t) + Kx(t) = 0, \tag{1.2}$$

where $M \in \mathscr{S}^n_+$, $C \in \mathscr{S}^n$ and $K \in \mathscr{S}^n_+$ represent the mass, damping and stiffness matrices, respectively. A critical observation noted in the survey paper [1] is that the solution of FEM can be obtained by solving the QEP (1.1). Unfortunately, the estimated solution of FEM does not match with the experimental results in some examples. For instance, Sheena, Unger and Zalmanovich [2] pointed out that the eigenvalues and the mode shape (the eigenvectors) obtained from the solution of the dynamic Eq. (1.2) with the corrected stiffness matrix are usually different from those under the corrected matrix. Various updating strategies have been proposed to deal with the improvement of FEM. Baruch [3,4] updated the stiffness matrix under the assumption that the mass matrix is set to be exact. Berman [5,6] further studied the inexact mass matrix and modified both the mass matrix and the stiffness matrix. Caesar [7] produced numerous approaches by optimizing different types of merit functions for (1.2). However, all the aforementioned approaches only changed one variable at a time. Wei [8–10] proposed several schemes to update the mass and stiffness matrices simultaneously and imposed some constraints on the model, such as mass orthogonality and the symmetry of the updated matrices. After that, many authors focus on the structure of the three quantities M, C, K in the model updating problem for (1.2). Friswell et al. [11] and Lancaster [12] directly solved the damping and stiffness matrices under the constraints that M is positive definite and C, K are symmetric, while Bai et al. [13] and Chu et al. [14] assumed that M, K are positive definite. Recently, the QEPs for vibrations of rotating machines or moving coordinate frames have been studied in [15,16], in which a *gyroscopic* system appears, that is,

$$M\ddot{x}(t) + (C + G)\dot{x}(t) + (K + N)x(t) = 0, \tag{1.3}$$

where the corresponding pencil $Q(\lambda)$ of (1.3) is defined by

$$Q(\lambda) := \lambda^2 M + \lambda(C + G) + (K + N) \tag{1.4}$$

and the skew-symmetric matrices $G, N \in \mathscr{K}^n$ represent the gyroscopic and circulatory matrices, respectively. More discussions on the gyroscopic system with applications can be found in [17–19]. Similar to [20], the given partially measured eigendatas in this paper are defined as follows:

$$\Lambda := \operatorname{diag}\left(\begin{bmatrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{bmatrix}, \ldots, \begin{bmatrix} \alpha_{s_c} & \beta_{s_c} \\ -\beta_{s_c} & \alpha_{s_c} \end{bmatrix}, \lambda_{2s_c+1}, \ldots, \lambda_k\right) \in \mathscr{R}^{k\times k}, \tag{1.5}$$

$$X := [x_{1R}, x_{1I}, \ldots, x_{s_cR}, x_{s_cI}, x_{2s_c+1}, \ldots, x_k] \in \mathscr{R}^{n\times k}, \tag{1.6}$$

with $1 \le k \le n$ and $\operatorname{rank}(X) = k$. Notice that the above eigendatas (Λ, X) satisfy the following *gyroscopic* structure [16]:

$$MX\Lambda^2 + (C + G)X\Lambda + (K + N)X = 0.$$

In this paper, we consider the *inverse damped gyroscopic eigenvalue problem* (IDGEP), which aims to find matrices closest to the given estimated matrices satisfying the given partially measured eigendatas. The IDGEP with the pencil $Q(\lambda)$ in (1.4) and (X, Λ) in (1.5–1.6) takes the form of

$$\begin{array}{ll} \min & \frac{1}{2}\|(M, C, K, G, N) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\ \text{s.t.} & (M, C, K, G, N) \in \mathscr{S}(X, \Lambda) \cap \Omega, \end{array} \tag{1.7}$$

where $(M_0, C_0, K_0, G_0, N_0)$ are the given estimated matrices and $\mathscr{S}(X, \Lambda)$ is defined by

$$\mathscr{S}(X, \Lambda) := \left\{(M, C, K, G, N) : MX\Lambda^2 + (C + G)X\Lambda + (K + N)X = 0\right\}. \tag{1.8}$$

Let X admit the QR factorization $X := Q\begin{bmatrix} R^T & 0^T \end{bmatrix}^T$. By renaming $M := Q^T M Q$, $C := Q^T C Q$, $K := Q^T K Q$, $G := Q^T G Q$, $N := Q^T N Q$, $M_0 := Q^T M_0 Q$, $C_0 := Q^T C_0 Q$, $K_0 := Q^T K_0 Q$, $G_0 := Q^T G_0 Q$ and $N_0 := Q^T N_0 Q$, the equality constraint in (1.8) becomes

$$M\begin{bmatrix} (R\Lambda^2)^T & 0^T \end{bmatrix}^T + (C+G)\begin{bmatrix} (R\Lambda)^T & 0^T \end{bmatrix}^T + (K+N)\begin{bmatrix} R^T & 0^T \end{bmatrix}^T = 0. \quad (1.9)$$

Similar to [21], we split each matrix into 2×2 blocks:

$$M := \begin{bmatrix} M_1 & M_2 \\ M_2^T & M_4 \end{bmatrix}, C := \begin{bmatrix} C_1 & C_2 \\ C_2^T & C_4 \end{bmatrix}, K := \begin{bmatrix} K_1 & K_2 \\ K_2^T & K_4 \end{bmatrix}, G := \begin{bmatrix} G_1 & G_2 \\ -G_2^T & G_4, \end{bmatrix},$$
$$N := \begin{bmatrix} N_1 & N_2 \\ -N_2^T & N_4, \end{bmatrix},$$

where $M_1, C_1, K_1 \in \mathscr{S}^k, M_4, C_4, K_4 \in \mathscr{S}^{n-k}, G_1, N_1 \in \mathscr{K}^k, G_4, N_4 \in \mathscr{K}^{n-k}, M_2, C_2, K_2, G_2, N_2 \in \mathscr{R}^{k\times(n-k)}$. Let $S := R\Lambda R^{-1}$. From (1.9) and the above notations, we rewrite (1.7) as follows:

$$(\mathscr{P}) \qquad \begin{array}{ll} \min & \Phi(M, C, K, G, N) \\ \text{s.t.} & \mathscr{H}_1(M, C, K, G, N) = 0, \\ & \mathscr{H}_2(M, C, K, G, N) = 0, \\ & (M, C, K, G, N) \in \Omega, \end{array}$$

where the objective function $\Phi : \Omega \to \mathscr{R}$ is defined by

$$\Phi(M, C, K, G, N) := \frac{1}{2}\|(M, C, K, G, N) - (M_0, C_0, K_0, G_0, N_0)\|^2$$

and $\mathscr{H}_i (i = 1, 2)$ are two linear operators as follows:

$$\begin{aligned} \mathscr{H}_1(M, C, K, G, N) &:= M_1 S^2 + (C_1 + G_1)S + (K_1 + N_1), \\ \mathscr{H}_2(M, C, K, G, N) &:= (S^2)^T M_2 + S^T(C_2 - G_2) + (K_2 - N_2). \end{aligned}$$

Let $\mathscr{H}_1^*$ and $\mathscr{H}_2^*$ denote the adjoint operators of $\mathscr{H}_1$ and $\mathscr{H}_2$, respectively. Given two arbitrary points $Y_1 \in \mathscr{R}^{k\times k}$, $Y_2 \in \mathscr{R}^{k\times(n-k)}$, we obtain

$$\mathscr{H}_1^*(Y_1) = \left(Y_1^1, Y_1^2, Y_1^3, Y_1^4, Y_1^5\right), \quad \mathscr{H}_2^*(Y_2) = \left(Y_2^1, Y_2^2, Y_2^3, Y_2^4, Y_2^5\right).$$

It is easy to verify that

$$\begin{aligned} &Y_1^1 = \mathscr{D}(\mathscr{T}(S^2, Y_1)), Y_1^2 = \mathscr{D}(\mathscr{T}(S, Y_1)), Y_1^3 = \mathscr{D}(\mathscr{T}(I_k, Y_1)), Y_1^4 = \mathscr{D}(\mathscr{T}^a(S, Y_1)), \\ &Y_1^5 = \mathscr{D}(\mathscr{T}^a(I_k, Y_1)), \\ &Y_2^1 = \mathscr{B}(S^2 Y_2), Y_2^2 = \mathscr{B}(SY_2), Y_2^3 = \mathscr{B}(Y_2), Y_2^4 = \mathscr{B}^a(SY_2), \\ &Y_2^5 = \mathscr{B}^a(Y_2), \end{aligned}$$

where

$$\mathscr{B}(X) := \frac{1}{2}\begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix}, \quad \mathscr{B}^a(X) := \frac{1}{2}\begin{bmatrix} 0 & -X \\ X^T & 0 \end{bmatrix}, \quad \mathscr{D}(X) := \frac{1}{2}\begin{bmatrix} X & 0 \\ 0 & 0 \end{bmatrix}$$

and $\mathscr{T}(X, Y) = XY + Y^T X^T$, $\mathscr{T}^a(X, Y) = Y^T X^T - XY$. In what follows, we refer to $(\mathscr{P})$ as the *inverse damped gyroscopic eigenvalue problem*. In this paper, we need the following assumption to hold:

The matrix Λ *is nonsingular*.

Remark 1.1 It follows from the above assumption that $(\mathscr{P})$ admits a strictly feasible solution. For the proof, see [21, Theorem 1]. In addition, the Lagrangian dual problem of $(\mathscr{P})$ is expressed by

$$(\mathscr{D}) \qquad \begin{array}{l} \min\ \Theta(Y_1, Y_2) \\ \text{s.t.} \quad (Y_1, Y_2) \in \mathscr{R}^{k\times k} \times \mathscr{R}^{k\times(n-k)}, \end{array}$$

where $\Theta : \mathscr{R}^{k\times k} \times \mathscr{R}^{k\times(n-k)} \to \mathscr{R}$ is defined as follows:

$$\begin{aligned} \Theta(Y_1, Y_2) := & \frac{1}{2}\left\| \Pi_\Omega \left((M_0, C_0, K_0, G_0, N_0) + \mathscr{H}_1^*(Y_1) + \mathscr{H}_2^*(Y_2)\right)\right\|^2 \\ & - \frac{1}{2}\|((M_0, C_0, K_0, G_0, N_0)\|^2. \end{aligned}$$

Recently, Xiao et al. [21] proposed an inexact smoothing Newton algorithm for the inverse damped gyroscopic eigenvalue problem (IDGEP), in which the Lagrangian dual problem $(\mathscr{D})$ was considered. The main purpose of this paper is to design an alternative algorithm for solving IDGEPs. The proposed algorithm is based on the framework of the augmented Lagrangian method and uses the accelerated proximal gradient method [22] as the solver of the corresponding subproblems. APG-type methods have been applied into various famous problems such as matrix completion problems [23], nearest correlation matrix problems [24] and signal processing problems [25]. These algorithms have an attractive iteration-complexity of $O(1/\sqrt{\varepsilon})$ to obtain an ε-optimal solution. The interested readers are referred to [26] and references therein. The theoretical attractive feature aforementioned is one of the important motivations to use the APG method as the core of our algorithm. From the computational experience in [23–25], APG-type methods may perform better than Newton-type algorithms, which is another motivation. Compared with the approach in [21], we directly solve the original problem $(\mathscr{P})$ in this paper based on the following observations: (a) Due to the analysis in Sect. 3, the subproblems of the prime problem $(\mathscr{P})$ are similar to (2.1), whereas the subproblems of $(\mathscr{D})$ have a different form. (b) These subproblems (see Remark 3.1) have a consistent structure and they are independent in each iteration. Consequently, we can solve these problems simultaneously. In the numerical experiments, we design a simple line search approach to solve these subproblems.

In order to illustrate the performance of the proposed algorithm, we compare it with the inexact smoothing Newton algorithm [21] and the interior-point method [16]. Unlike other algorithms for the IDGEP, our algorithm ALA has the following two main features:

(a) The whole sequence $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ generated by ALA converges to the unique solution of the IDGEP.
(b) Without the additional regularity condition (such as the constraint nondegeneracy for the inexact smoothing Newton algorithm in [21]), the proposed algorithm only needs at most $O(\log(\varepsilon^{-1}))$ outer iterations and at most $O(\varepsilon^{-1})$ APG calls to obtain an ε-feasible and ε-optimal solution of the IDGEP.

The contribution of this paper is to establish an iteration-complexity bound for the proposed algorithm to solve the inverse damped gyroscopic eigenvalue problem, while in [21] there is no discussion on the iteration-complexity of the inexact smoothing Newton algorithm. We emphasize this point because complexity analysis is an indispensable ingredient for algorithm design. See [27] and references therein.

The rest of the paper is organized as follows. In Sect. 2, we review some technical results on the accelerated proximal gradient method. Sect. 3 is devoted to designing an alternative approach for the IDGEP by combining the augmented Lagrangian method with the APG strategy. The corresponding convergence results and the iteration-complexity bound are established in Sect. 4. In Sect. 5, numerical results are reported. Finally, we present some concluding remarks.

2 Technical results

In this section, we first review the accelerated proximal gradient method and then present some technical results used in the sequel. Consider a problem of minimizing a structured function Υ over a set Ξ:

$$\min_{W\in\Xi} \Upsilon(W) := P(W) + F(W), \tag{2.1}$$

where Ξ is a finite dimensional Hilbert space, $P : \Xi \to (-\infty, +\infty)$ and $F : \Xi \to (-\infty, +\infty]$ are proper, lower semi-continuous and convex functions such that dom P (the effective domain of P) is closed. Moreover, the gradient ∇F is Lipschitz continuous on dom P for some $L_F > 0$, that is,

$$\|\nabla F(W_1) - \nabla F(W_2)\|_F \le L_F \|W_1 - W_2\|_F.$$

The class of convex functions satisfying the above relation is denoted by $\mathscr{C}^{1,1}_{L_F}$.

Next, we present an important property of a function in $\mathscr{C}^{1,1}_{L_F}$, see [27, Lemma 1.2.3] for the proof.

Lemma 2.1 *The conclusion holds for any $F \in \mathscr{C}^{1,1}_{L_F}$ and $W_1, W_2 \in$ dom F:*

$$|F(W_1) - F(W_2) - \langle \nabla F(W_2), W_1 - W_2\rangle| \le \frac{L_F}{2}\|W_1 - W_2\|_F^2.$$

Remark 2.2 From Lemma 2.1, let us denote

$$Q(W,Z) = P(W) + F(Z) + \langle \nabla F(Z), W - Z\rangle + \frac{L_F}{2}\|W - Z\|_F^2, \; Z \in \operatorname{dom} P$$

and

$$\bar{W}(Z) = \operatorname{argmin}\{Q(W,Z) : W \in \mathscr{S}\}, \tag{2.2}$$

where $\mathscr{S} \subset \Xi$ is a convex set such that the optimal solution W^* of $\Upsilon(W)$ lies in $\mathscr{S}$. Then, we obtain $Q(W,Z) \geq \Upsilon(W)$. In addition, $Q(W,Z)$ is a strongly convex function with respect to W.

The accelerated proximal gradient method is displayed in the Algorithm 1 below, where APG-stop(s) denotes the flag of the termination criterion for Algorithm 1, which only depends on the counter s. The details of APG-stop(s) are specified in the Algorithm 2.

Algorithm 1 Accelerate Proximal Gradient Method (APG)

1: Choose $W_1^0 = W_2^1 = W^0 \in \operatorname{dom} P$, $t_1 = 1$ and set $s = 0$.
2: **repeat**
3: Set $s = s + 1$.
4: Set $W_1^s = \bar{W}(W_2^s)$, where $\bar{W}(\cdot)$ is defined as (2.2).
5: Set $t_{s+1} = \frac{1+\sqrt{1+4t_s^2}}{2}$.
6: Set $W_2^{s+1} = W_1^s + \frac{t_s - 1}{t_{s+1}}(W_1^s - W_1^{s-1})$.
7: **until** APG-stop(s) is **true**.
8: **return** W_1^s.

We presents $O(\sqrt{L_F/\varepsilon})$ iteration-complexity for Algorithm 1. The interested readers are referred to Corollary 3 in [22] and Theorem 4.4 in [25] for the proof.

Lemma 2.3 *Let $\{W_1^s\}$ be the sequence generated by Algorithm 1, then for any $s \geq 1$, we have*

$$\Upsilon(W_1^s) - \min_{W \in \Xi} \Upsilon(W) \leq \varepsilon,$$

whenever $s \geq \sqrt{\frac{2L_F}{\varepsilon}}\|W^* - W^0\|_F - 1$ and $W^* := \operatorname{argmin}_{W \in \Xi}\Upsilon(W)$.

To end this section, a fundamental connection between the normal cone and optimality conditions is provided, which comes from [28, Theorem 6.12].

Lemma 2.4 *Consider a problem of minimizing a differentiable function f_0 over a set C. A necessary condition for $\bar{x}$ to be locally optimal is*

$$0 \in \nabla f_0(\bar{x}) + \mathscr{N}_C(\bar{x}),$$

where $\mathcal{N}_C(\bar{x})$ denotes the normal cone of C at $\bar{x}$. When C and f_0 are both convex, the above condition is sufficient for $\bar{x}$ to be globally optimal and can be written also in the form

$$\langle \nabla f_0(\bar{x}), x - \bar{x}\rangle \geq 0 \quad \textit{for all } x \in C.$$

3 Augmented Lagrangian method based on the APG strategy for the IDGEP

In this section, we design an augmented Lagrangian method for the inverse damped gyroscopic eigenvalue problem. The augmented Lagrangian function of the IDGEP is defined as follows:

$$\begin{aligned}
&\mathcal{L}(M, C, K, G, N, \Gamma_1, \Gamma_2, \rho) := \rho\Phi(M, C, K, G, N) \\
&\quad + \frac{1}{2}\sum_{i=1}^{2}\Big(\|\mathcal{H}_i(M, C, K, G, N) - \rho\Gamma_i\|_F^2 - \|\rho\Gamma_i\|_F^2\Big).
\end{aligned}$$

We solve the following sequence of subproblems inexactly,

$$\begin{aligned}
&\min\ \Upsilon^\nu(M, C, K, G, N) := \mathcal{L}(M, C, K, G, N, \Gamma_1^\nu, \Gamma_2^\nu, \rho_\nu) + \frac{\rho_\nu^2}{2}\sum_{i=1}^{2}\|\Gamma_i^\nu\|_F^2 \\
&\text{s.t.}\quad (M, C, K, G, N) \in \Omega.
\end{aligned} \tag{3.1}$$

For simplicity, we rewrite (3.1) as follows:

$$(\mathcal{P}^\nu) \qquad \begin{aligned}
&\min\ \Upsilon^\nu(M, C, K, G, N) := P^\nu(M, C, K, G, N) + F^\nu(M, C, K, G, N) \\
&\text{s.t.}\quad (M, C, K, G, N) \in \Omega,
\end{aligned}$$

where $P^\nu : \Omega \to \mathcal{R}$ and $F^\nu : \Omega \to \mathcal{R}$ are defined by

$$P^\nu(M, C, K, G, N) := \rho_\nu\Phi(M, C, K, G, N), \tag{3.2}$$

$$F^\nu(M, C, K, G, N) := \frac{1}{2}\sum_{i=1}^{2}\|\mathcal{H}_i(M, C, K, G, N) - \rho_\nu\Gamma_i^\nu\|_F^2. \tag{3.3}$$

Notice that if we set

$$\begin{aligned}
W &:= (M, C, K, G, N),\ \ \Xi := \mathcal{S}^n \times \mathcal{S}^n \times \mathcal{S}^n \times \mathcal{R}^{n\times n} \times \mathcal{R}^{n\times n}, \\
P(W) &:= P^\nu(M, C, K, G, N) + \delta_\Omega(M, C, K, G, N), \\
F(W) &:= F^\nu(M, C, K, G, N),
\end{aligned}$$

$(\mathcal{P}^\nu)$ has the form similar to (2.1).

In the Algorithm 2, the augmented Lagrangian method based on the APG strategy for the inverse damped gyroscopic eigenvalue problem $(\mathcal{P})$ is presented. In each iteration, we call Algorithm 1 to solve $(\mathscr{P}^\nu)$ inexactly. Similar to Algorithm 1, ITER-stop(s) and GARD-stop(x) represent the flags of the stopping criterions that only depends on the current iteration counter s and the iterative point x, respectively. In addition, APG-stop(s) holds if one of the flags ITER-stop(s) and GARD-stop(x) is true. The termination criterion for Algorithm 2 is denoted by ALA-stop, which is specified in Sect. 5.2.

Algorithm 2 Augmented Lagrangian Method Based on the APG Strategy for the IDGEP (ALA)

1: Given the initial point $(M^0, C^0, K^0, G^0, N^0) \in \Omega_0$ satisfying $\mathscr{H}_1(M^0, C^0, K^0, G^0, N^0) = 0$ and $\mathscr{H}_2(M^0, C^0, K^0, G^0, N^0) = 0$.
2: Set $\Gamma_1^1 := 0$, $\Gamma_2^1 := 0$ and $\nu := 0$.
3: **repeat**
4: Set $\nu := \nu + 1$ and $\mathscr{S}_\nu$ is defined as (3.5).
5: Call Algorithm 1 to solve $(\mathscr{P}^\nu)$ and generate the sequence $\{(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s})\}$, in which we set $(M^{\nu-1}, C^{\nu-1}, K^{\nu-1}, G^{\nu-1}, N^{\nu-1})$ as the initial point and the convex set $\mathscr{S} := \mathscr{S}_\nu$ in (2.2). Set ITER-stop(s) $:= \{s \geq s_{\max}^\nu$, where $s_{\max}^\nu$ is defined as (3.13)$\}$, GARD-stop$(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s}) := \{\|(TM_1^{\nu,s}, TC_1^{\nu,s}, TK_1^{\nu,s}, TG_1^{\nu,s}, TN_1^{\nu,s})\| \leq \tau^\nu\}$ and

$$
\begin{aligned}
TM_1^{\nu,s} &:= \Pi_{\mathscr{S}_+^n}\left(\rho_\nu M_0 + L_F M_1^{\nu,s} - \nabla_M F^\nu(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s})\right) - (\rho_\nu + L_F)M_1^{\nu,s},\\
TC_1^{\nu,s} &:= \rho_\nu(C_1^{\nu,s} - C_0) + \nabla_C F^\nu(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s}),\\
TK_1^{\nu,s} &:= \Pi_{\mathscr{S}_+^n}\left(\rho_\nu K_0 + L_F K_1^{\nu,s} - \nabla_K F^\nu(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s})\right) - (\rho_\nu + L_F)K_1^{\nu,s},\\
TG_1^{\nu,s} &:= \rho_\nu(G_1^{\nu,s} - G_0) + \nabla_G F^\nu(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s}),\\
TN_1^{\nu,s} &:= \rho_\nu(N_1^{\nu,s} - N_0) + \nabla_N F^\nu(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s}).
\end{aligned}
$$

The entries of $\nabla F^\nu(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s})$ are defined as (3.8).
6: Set $(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) := (M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s})$.
7: Set $\Gamma_i^{\nu+1} := \Gamma_i^\nu - \rho_\nu^{-1}\mathscr{H}_i(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)$ $(i = 1, 2)$.
8: **until** ALA-stop is **true**.
9: **return** $(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)$.

3.1 The accelerate gradient method for the subproblem $(\mathscr{P}^\nu)$

From Step 4 of Algorithm 1, for the given point $(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}) \in \Omega_0$, we need to solve the following problem

$$
\begin{aligned}
\min \quad & P^\nu(M, C, K, G, N) + \langle \nabla F^\nu(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}), (M, C, K, G, N) - (\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N})\rangle \\
& + \frac{L_F}{2}\left(\|M - M_0\|_F^2 + \|C - C_0\|_F^2 + \|K - K_0\|_F^2 + \|G - G_0\|_F^2 + \|N - N_0\|_F^2\right)\\
\text{s.t.} \quad & (M, C, K, G, N) \in \mathscr{S}_\nu,
\end{aligned}
\tag{3.4}
$$

where $L_F := \|\hat{E}^T\hat{E}\|_F$, $\hat{E} := [S^2 \quad S \quad I_k \quad -S \quad -I_k]$ (see Appendix 1). The convex set $\mathscr{S}_\nu$ is defined by

$$\mathscr{S}_\nu := \left\{(M, C, K, G, N) \in \Omega : \begin{array}{c} \|M - M_0\|_F \le \eta^\nu, \|C - C_0\|_F \le \eta^\nu, \|K - K_0\|_F \le \eta^\nu, \\ \|G - G_0\|_F \le \eta^\nu, \|N - N_0\|_F \le \eta^\nu \end{array}\right\}, \tag{3.5}$$

$$\eta^\nu := \sqrt{2\rho_\nu^{-1}\Upsilon^\nu(M^0, C^0, K^0, G^0, N^0)} \tag{3.6}$$

and ∇F^ν denotes the gradient of F^ν, that is,

$$\nabla F^\nu\left(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}\right) = \sum_{i=1}^{2} \mathscr{H}_i^*\left(\mathscr{H}_i\left(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}\right) - \rho_\nu \Gamma_i^\nu\right). \tag{3.7}$$

It follows from (3.7) that the entries of ∇F^ν are defined as follows:

$$\begin{aligned}
\nabla_M F^\nu\left(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}\right) &= \mathscr{D}\left(\mathscr{T}\left(S^2, \tilde{H}_1^\nu\right)\right) + \mathscr{B}\left(S^2\tilde{H}_2^\nu\right),\\
\nabla_C F^\nu\left(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}\right) &= \mathscr{D}\left(\mathscr{T}\left(S, \tilde{H}_1^\nu\right)\right) + \mathscr{B}\left(S\tilde{H}_2^\nu\right),\\
\nabla_K F^\nu\left(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}\right) &= \mathscr{D}\left(\mathscr{T}\left(I_k, \tilde{H}_1^\nu\right)\right) + \mathscr{B}\left(\tilde{H}_2^\nu\right),\\
\nabla_G F^\nu\left(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}\right) &= \mathscr{D}\left(\mathscr{T}^a\left(S, \tilde{H}_1^\nu\right)\right) + \mathscr{B}^a\left(S\tilde{H}_2^\nu\right),\\
\nabla_N F^\nu\left(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}\right) &= \mathscr{D}\left(\mathscr{T}^a\left(I_k, \tilde{H}_1^\nu\right)\right) + \mathscr{B}^a\left(\tilde{H}_2^\nu\right),
\end{aligned} \tag{3.8}$$

where $\tilde{H}_i^\nu := \mathscr{H}_i(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}) - \rho_\nu\Gamma_i^\nu$, $(i = 1, 2)$.

Remark 3.1 Notice that the objective function in (3.4) has a separable structure. Hence, (3.4) can be divided into the five independent subproblems as follows:

$$(\mathscr{P}_M^\nu)\ \begin{cases} \min\ \frac{\rho_\nu}{2}\|M - M_0\|_F^2 + \langle \nabla_M F^\nu(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}), M - \tilde{M}\rangle + \frac{L_F}{2}\|M - \tilde{M}\|_F^2 \\ \text{s.t.}\quad M \in \mathscr{R}(M_0, \eta^\nu) := \{M \in \mathscr{S}_+^n : \|M - M_0\|_F \le \eta^\nu\}, \end{cases}$$

$$(\mathscr{P}_C^\nu)\ \begin{cases} \min\ \frac{\rho_\nu}{2}\|C - C_0\|_F^2 + \langle \nabla_C F^\nu(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}), C - \tilde{C}\rangle + \frac{L_F}{2}\|C - \tilde{C}\|_F^2 \\ \text{s.t.}\quad C \in \mathscr{R}(C_0, \eta^\nu) := \{C \in \mathscr{S}^n : \|C - C_0\|_F \le \eta^\nu\}, \end{cases}$$

$$(\mathscr{P}_K^\nu)\ \begin{cases} \min\ \frac{\rho_\nu}{2}\|K - K_0\|_F^2 + \langle \nabla_K F^\nu(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}), K - \tilde{K}\rangle + \frac{L_F}{2}\|K - \tilde{K}\|_F^2 \\ \text{s.t.}\quad K \in \mathscr{R}(K_0, \eta^\nu) := \{K \in \mathscr{S}_+^n : \|K - K_0\|_F \le \eta^\nu\}, \end{cases}$$

$$(\mathscr{P}_G^\nu)\ \begin{cases} \min\ \frac{\rho_\nu}{2}\|G - G_0\|_F^2 + \langle \nabla_G F^\nu(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}), G - \tilde{G}\rangle + \frac{L_F}{2}\|G - \tilde{G}\|_F^2 \\ \text{s.t.}\quad G \in \mathscr{R}(G_0, \eta^\nu) := \{G \in \mathscr{K}^n : \|G - G_0\|_F \le \eta^\nu\}, \end{cases}$$

$$(\mathscr{P}_N^\nu)\ \begin{cases} \min\ \frac{\rho_\nu}{2}\|N - N_0\|_F^2 + \langle \nabla_N F^\nu(\tilde{M}, \tilde{C}, \tilde{K}, \tilde{G}, \tilde{N}), N - \tilde{N}\rangle + \frac{L_F}{2}\|N - \tilde{N}\|_F^2 \\ \text{s.t.}\quad N \in \mathscr{R}(N_0, \eta^\nu) := \{N \in \mathscr{K}^n : \|N - N_0\|_F \le \eta^\nu\}. \end{cases}$$

More discussions on how to solve these subproblems are presented in Sect. 5.1.

To conclude this subsection, the following lemma shows that the minimizer of $(\mathscr{P}^\nu)$ lies in the pre-defined set $\mathscr{S}_\nu$.

Lemma 3.2 *Let $(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)$ be an optimal solution of $(\mathscr{P}^\nu)$. Then, we obtain*

$$\mathscr{S}_\nu \cap argmin\{\Upsilon^\nu(M, C, K, G, N) : (M, C, K, G, N) \in \Omega\} \neq \emptyset,$$

where $\mathscr{S}_\nu$ is defined as (3.5).

Proof Since $(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)$ is an optimal solution of $(\mathscr{P}^\nu)$, we have

$$P^\nu(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu) \leq \Upsilon^\nu(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu) \leq \Upsilon^\nu(M^0, C^0, K^0, G^0, N^0),$$

which yields that

$$\begin{aligned} &\|(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu) - (M_0, C_0, K_0, G_0, N_0)\| \\ &\leq \sqrt{2\rho_\nu^{-1}\Upsilon^\nu\left(M^0, C^0, K^0, G^0, N^0\right)} = \eta^\nu. \end{aligned} \tag{3.9}$$

Hence, we deduce that $\mathscr{S}_\nu \cap \text{argmin}\{\Upsilon^\nu(M, C, K, G, N) : (M, C, K, G, N) \in \Omega\} \neq \emptyset$. □

3.2 The discussions on the stopping criterions ITER-stop and GARD-stop

Next, we analyze the stopping criterion APG-stop at Algorithm 2. Let $\{(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s})\}$ be the sequence generated by calling Algorithm 1 to solve $(\mathscr{P}^\nu)$. For simplicity, the function d^ν is defined as follows:

$$d^\nu(M, C, K, G, N) := \|(M, C, K, G, N) - \left(M^{\nu-1}, C^{\nu-1}, K^{\nu-1}, G^{\nu-1}, N^{\nu-1}\right)\|^2. \tag{3.10}$$

Let $(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)$ be an optimal solution of $(\mathscr{P}^\nu)$. From Lemma 2.3 and (3.10), when

$$s \geq \sqrt{\frac{2L_F d^\nu(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)}{\varepsilon^\nu}} - 1, \tag{3.11}$$

the following inequality holds at the point $(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s})$:

$$\begin{aligned} &\Upsilon^\nu(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s}) \\ &\leq \inf\left\{\Upsilon^\nu(M, C, K, G, N) : (M, C, K, G, N) \in \Omega\right\} + \varepsilon^\nu. \end{aligned} \tag{3.12}$$

In what follows, let us denote

$$s_{\max}^\nu = \sqrt{\frac{2L_F \xi^\nu}{\varepsilon^\nu}}, \tag{3.13}$$

where $\xi^\nu := d^1(M_0, C_0, K_0, G_0, N_0) + d^\nu(M_0, C_0, K_0, G_0, N_0) + \rho_\nu \sum_{i=1}^2 \|\Gamma_i^\nu\|_F^2$. It is easy to verify that (see Appendix 2)

$$s_{\max}^\nu \ge \sqrt{\frac{2L_F d^\nu(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)}{\varepsilon^\nu}}.$$

From (3.12) and (3.13), the point $(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s})$ is an ε^ν-optimal solution of $(\mathscr{P}^\nu)$, provided that the ITER-stop of Algorithm 2 is true.

Finally, we analyze the stopping criterion GARD-stop to end this section. From Lemma 2.4, we obtain

$$0 \in \nabla\Upsilon^\nu(M, C, K, G, N) + \mathscr{N}_\Omega(M, C, K, G, N), \tag{3.14}$$

where the entries of $\nabla\Upsilon^\nu(M, C, K, G, N)$ are expressed by

$$\begin{aligned}
\nabla_M\Upsilon^\nu(M, C, K, G, N) &= \rho_\nu(M - M_0) + \nabla_M F^\nu(M, C, K, G, N),\\
\nabla_C\Upsilon^\nu(M, C, K, G, N) &= \rho_\nu(C - C_0) + \nabla_C F^\nu(M, C, K, G, N),\\
\nabla_K\Upsilon^\nu(M, C, K, G, N) &= \rho_\nu(K - K_0) + \nabla_K F^\nu(M, C, K, G, N),\\
\nabla_G\Upsilon^\nu(M, C, K, G, N) &= \rho_\nu(G - G_0) + \nabla_G F^\nu(M, C, K, G, N),\\
\nabla_N\Upsilon^\nu(M, C, K, G, N) &= \rho_\nu(N - N_0) + \nabla_N F^\nu(M, C, K, G, N).
\end{aligned} \tag{3.15}$$

Combining (3.14, 3.15) with the GARD-stop, the point $(M_1^{\nu,s}, C_1^{\nu,s}, K_1^{\nu,s}, G_1^{\nu,s}, N_1^{\nu,s})$ obtained by Algorithm 2 satisfies the first-order optimality condition of $(\mathscr{P}^\nu)$ if we set $\tau^\nu \equiv 0$.

4 Convergence results and the iteration-complexity analysis

In this section, we present the convergence results and the iteration-complexity for Algorithm 2. Let $(\bar{M}^\nu, \bar{C}^\nu, \bar{K}^\nu, \bar{G}^\nu, \bar{N}^\nu)$ be an ε^ν-optimal solution of $(\mathscr{P}^\nu)$, that is,

$$\Upsilon^\nu(\bar{M}^\nu, \bar{C}^\nu, \bar{K}^\nu, \bar{G}^\nu, \bar{N}^\nu) \le \min\{\Upsilon^\nu(M, C, K, G, N) : (M, C, K, G, N) \in \Omega\} + \varepsilon^\nu.$$

Then, we have (see Appendix 3)

$$\begin{aligned}
\|\nabla_C F^\nu(\bar{M}^\nu, \bar{C}^\nu, \bar{K}^\nu, \bar{G}^\nu, \bar{N}^\nu)\|_F^2 \ &+ \|\nabla_G F^\nu(\bar{M}^\nu, \bar{C}^\nu, \bar{K}^\nu, \bar{G}^\nu, \bar{N}^\nu)\|_F^2\\
&\le 2(\rho_\nu + L_F)\varepsilon^\nu.
\end{aligned} \tag{4.1}$$

From Step 5 of Algorithm 2, if ITER-stop(s) is true, the point $(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)$ is an ε^ν-optimal solution of $(\mathscr{P}^\nu)$. Let us denote $H_i^\nu := \mathscr{H}_i(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) - \rho_\nu\Gamma_i^\nu$, $(i = 1, 2)$. The inequality (4.1) implies that

$$\|\mathscr{D}(\mathscr{T}(S, H_1^\nu)) + \mathscr{B}(SH_2^\nu)\|_F^2 + \|\mathscr{D}(\mathscr{T}^a(S, H_1^\nu)) + \mathscr{B}^a(SH_2^\nu)\|_F^2 \le 2(\rho_\nu + L_F)\varepsilon^\nu. \tag{4.2}$$

Similarly, if GARD-stop$(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)$ is true, we obtain

$$\begin{aligned}
&\|\mathscr{D}(\mathscr{T}(S, H_1^\nu)) + \mathscr{B}(SH_2^\nu)\|_F^2 + \|\mathscr{D}(\mathscr{T}^a(S, H_1^\nu)) + \mathscr{B}^a(SH_2^\nu)\|_F^2 \\
&\quad \le (\tau^\nu)^2 + \rho_\nu^2(\|C_0 - C^\nu\|_F^2 + \|G_0 - G^\nu\|_F^2).
\end{aligned} \tag{4.3}$$

We begin by establishing the bounds on the sequence of the Lagrangian multipliers $\Gamma_i^\nu (i = 1, 2)$.

Lemma 4.1 *The following conclusion holds:*

$$\begin{aligned}
&\sum_{i=1}^{2} \|\Gamma_i^{\nu+1}\|_F^2 \\
&\quad \le \max\left\{2\rho_\nu^{-2}\eta_S(\rho_\nu + L_F)\varepsilon^\nu, \frac{1}{1 - 2\eta_S\rho_{\nu_0}}\eta_S(2d + (\tau^{\nu_0}/\rho_{\nu_0})^2) + T^{\nu_0}\right\},
\end{aligned} \tag{4.4}$$

where ν_0 is a positive integer such that $2\eta_S\rho_{\nu_0} < 1$, $\{\tau^\nu\}$ and $\{\rho_\nu\}$ are two decreasing sequences satisfying the condition:

$$\frac{\tau^{\nu+1}}{\rho_{\nu+1}} \le \frac{\tau^\nu}{\rho_\nu}$$

and the sequence $\{T^{\nu+1}\}$ satisfy $T^{\nu+1} := \eta_S(\tau^\nu/\rho_\nu)^2 + 2\eta_S d + (2\eta_S\rho_\nu)T^\nu$, $T^1 := 0$, $\eta_S := \|S^{-1}\|_F^2$, $d := d^1(M_0, C_0, K_0, G_0, N_0)$ and $d^1(M_0, C_0, K_0, G_0, N_0)$ is defined as (3.10).

Proof From (4.2) and the definitions of $\mathscr{D}, \mathscr{B}, \mathscr{B}^a, \mathscr{T}, \mathscr{T}^a$, we have

$$\begin{aligned}
&\left\| \frac{1}{2} \begin{bmatrix} SH_1^\nu + (H_1^\nu)^T S^T & SH_2^\nu \\ (H_2^\nu)^T S^T & 0 \end{bmatrix} \right\|_F^2 + \left\| \frac{1}{2} \begin{bmatrix} (H_1^\nu)^T S^T - SH_1^\nu & -SH_2^\nu \\ (H_2^\nu)^T S^T & 0 \end{bmatrix} \right\|_F^2 \\
&\quad \le 2(\rho_\nu + L_F)\varepsilon^\nu,
\end{aligned}$$

which implies that

$$\|SH_1^\nu\|_F^2 + \|SH_2^\nu\|_F^2 \le 2(\rho_\nu + L_F)\varepsilon^\nu \;\Rightarrow\; \sum_{i=1}^{2} \|H_i^\nu\|_F^2 \le 2\eta_S(\rho_\nu + L_F)\varepsilon^\nu. \tag{4.5}$$

According to the inequality (4.3), we obtain

$$\sum_{i=1}^{2} \|H_i^\nu\|_F^2 \le \eta_S \left((\tau^\nu)^2 + \rho_\nu^2(\|C_0 - C^\nu\|_F^2 + \|G_0 - G^\nu\|_F^2) \right). \tag{4.6}$$

It follows from the definition of $\mathscr{S}_\nu$ and (3.6) that $\|C_0 - C^\nu\|_F^2 + \|G_0 - G^\nu\|_F^2 \le 2(\eta^\nu)^2$. In addition, the inequality (4.6) can be rewritten as follows:

$$\sum_{i=1}^{2} \|H_i^\nu\|_F^2 \le \eta_S(\tau^\nu)^2 + 2\eta_S\rho_\nu^2 d + 2\eta_S\rho_\nu^3 \sum_{i=1}^{2} \|\Gamma_i^\nu\|_F^2. \tag{4.7}$$

It follows from Step 7 of Algorithm 2, (4.5) and (4.7) that

$$\sum_{i=1}^{2} \|\Gamma_i^{\nu+1}\|_F^2 \le 2\rho_\nu^{-2}\eta_S(\rho_\nu + L_F)\varepsilon^\nu, \tag{4.8}$$

$$\sum_{i=1}^{2} \|\Gamma_i^{\nu+1}\|_F^2 \le \eta_S(\tau^\nu/\rho_\nu)^2 + 2\eta_S d + 2\eta_S\rho_\nu \sum_{i=1}^{2} \|\Gamma_i^\nu\|_F^2. \tag{4.9}$$

For simplicity, we define the following notations:

$$\Gamma^\nu := (\Gamma_1^\nu,\ \Gamma_2^\nu),\ J^\nu := \eta_S(\tau^\nu/\rho_\nu)^2 + 2\eta_S d,\ \mu_\nu := 2\eta_S\rho_\nu.$$

The inequality (4.9) becomes

$$\|\Gamma^{\nu+1}\|_F^2 \le J^\nu + \mu_\nu\|\Gamma^\nu\|_F^2,\ \forall \nu \ge 1. \tag{4.10}$$

Now, let us prove the assertion that *the term* $\|\Gamma^\nu\|_F^2$ *is bounded for any* $\nu \ge 1$. Denote the sequence $\{T^{\nu+1}\}$ satisfying $T^{\nu+1} := J^\nu + \mu_\nu T^\nu$ and $T^1 := 0$. It follows from the inequality (4.10) that

$$\|\Gamma^\nu\|_F^2 \le T^\nu, \quad \forall \nu \ge 1. \tag{4.11}$$

Since $\{\rho_\nu\}$ is a decreasing sequence, there exists a positive integer ν_0 such that $\mu_{\nu_0} < 1$ and

$$\|\Gamma^{\nu_0+1}\|_F^2 \le J^{\nu_0} + \mu_{\nu_0}\|\Gamma^{\nu_0}\|_F^2 \le J^{\nu_0} + \mu_{\nu_0}T^{\nu_0} \le J^{\nu_0} + T^{\nu_0}.$$

Combining the above inequality with (4.10), we obtain

$$\|\Gamma^{\nu_0+2}\|_F^2 \le J^{\nu_0+1} + \mu_{\nu_0+1}\|\Gamma^{\nu_0+1}\|_F^2 \le J^{\nu_0} + \mu_{\nu_0}(J^{\nu_0} + T^{\nu_0}) \le (1 + \mu_{\nu_0})J^{\nu_0} + T^{\nu_0},$$

where the second inequality follows from the condition (4.5) and $\mu_{\nu_0} < 1$. Similarly, for any integer $m \ge 1$, we have

$$\|\Gamma^{\nu_0+m}\|_F^2 \le J^{\nu_0}\sum_{i=0}^{m-1}\mu_{\nu_0}^i + T^{\nu_0} \le \frac{1}{1-\mu_{\nu_0}}J^{\nu_0} + T^{\nu_0}.$$

The above inequality and (4.11) tell us that

$$\|\Gamma^\nu\|_F^2 \le \frac{1}{1-\mu_{\nu_0}}J^{\nu_0} + T^{\nu_0} \quad \Rightarrow$$

$$\sum_{i=1}^{2}\|\Gamma_i^{\nu+1}\|_F^2 \le \frac{1}{1-2\eta_S\rho_{\nu_0}}\eta_S(2d + (\tau^{\nu_0}/\rho_{\nu_0})^2) + T^{\nu_0}. \tag{4.12}$$

Hence, the conclusion (4.4) is obtained from (4.8) with (4.12). □

The following lemma shows that the sequences $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ and $\{(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)\}$ are bounded under some mild conditions for the parameters τ^ν, ρ_ν and ε^ν.

Lemma 4.2 *Let $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ be the sequence generated by Algorithm 2 with the parameters $\rho_\nu, \varepsilon^\nu, \tau^\nu$ satisfying the following conditions:*

(i) $\rho_\nu \searrow 0, \tau^\nu \searrow 0$ *such that* $(\tau^\nu/\rho_\nu) \searrow 0$.
(ii) $\varepsilon^\nu \searrow 0$ *and there exists a positive scalar* ζ *such that* $\varepsilon^\nu/(\rho_\nu^2) \leq \zeta$ *for all* $\nu \geq 1$.

The sequence of optimal solutions for $(\mathscr{P}^\nu)$ is denoted by $\{(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)\}$. Then, the sequences $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ and $\{(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)\}$ are bounded.

Proof From the definition of $\mathscr{S}_\nu$ and (3.9), we only need to verify that the sequence $\{\eta^\nu\}$ is bounded. Combining Lemma 4.1 with the definition of η^ν, we obtain

$$\begin{aligned}\eta^\nu &= \sqrt{\|(M^0, C^0, K^0, G^0, N^0) - (M_0, C_0, K_0, G_0, N_0)\|^2 + \rho_\nu \sum_{i=1}^{2} \|\Gamma_i^\nu\|_F^2} \\ &\leq \sqrt{d + \rho_1 \max\left\{2\eta_S(\rho_1 + L_F)\zeta, \frac{1}{1 - 2\eta_S\rho_{\nu_0}}\eta_S(2d + (\tau^{\nu_0}/\rho_{\nu_0})^2) + T^{\nu_0}\right\}} \\ &=: B_\eta,\end{aligned} \tag{4.13}$$

where $d := d^1(M_0, C_0, K_0, G_0, N_0)$ and $d^1(M_0, C_0, K_0, G_0, N_0)$ is defined as (3.10). It follows from (4.13) that $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ and $\{(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)\}$ are two bounded sequences. □

After these preparations, the convergence result for Algorithm 2 is established in the following theorem.

Theorem 4.1 *Let $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ be the sequence generated by Algorithm 2 with the parameters $\rho_\nu, \varepsilon^\nu, \tau^\nu$ satisfying the conditions mentioned in Lemma 4.2. Then, any limit point $(\bar{M}, \bar{C}, \bar{K}, \bar{G}, \bar{N})$ of the sequence $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ is an optimal solution of $(\mathscr{P})$.*

Proof From Lemma 4.2, the sequence of $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ is bounded. Hence, there exists an index set I such that

$$\lim_{\nu \in I}(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) = (\bar{M}, \bar{C}, \bar{K}, \bar{G}, \bar{N}). \tag{4.14}$$

In addition, when the index $\nu \in I$ converges to $+\infty$, we obtain

$$\lim_{\nu \in I} \tau^\nu = 0, \quad \lim_{\nu \in I} \rho_\nu = 0. \tag{4.15}$$

The inequalities (4.6) and (4.7) tell us that

$$\lim_{\nu\in I}\sum_{i=1}^{2}\|H_i^\nu\|_F^2 \le \lim_{\nu\in I}\|S^{-1}\|_F^2\left((\tau^\nu)^2+\rho_\nu^2\left(\|C_0-C^\nu\|_F^2+\|G_0-G^\nu\|_F^2\right)\right),$$

$$\lim_{\nu\in I}\sum_{i=1}^{2}\|H_i^\nu\|_F^2 \le \lim_{\nu\in I}\|S^{-1}\|_F^2(\tau^\nu)^2+2\|S^{-1}\|_F^2\rho_\nu^2 d+2\|S^{-1}\|_F^2\rho_\nu^3\sum_{i=1}^{2}\|\Gamma_i^\nu\|_F^2.$$

It follows from (4.14–4.15) and Lemma 4.1 that

$$0=\lim_{\nu\in I}H_1^\nu=\mathscr{H}_1(\bar{M},\bar{C},\bar{K},\bar{G},\bar{N}),\quad 0=\lim_{\nu\in I}H_2^\nu=\mathscr{H}_2(\bar{M},\bar{C},\bar{K},\bar{G},\bar{N}),$$

which means that $(\bar{M},\bar{C},\bar{K},\bar{G},\bar{N})$ is a feasible solution of $(\mathscr{P})$.

In the sequel, we show that $(\bar{M},\bar{C},\bar{K},\bar{G},\bar{N})$ is indeed an optimal solution of $(\mathscr{P})$. The proof includes the following two cases:

(A): $|I_I|=+\infty$, where $I_I\subseteq I$ denotes the index set that Algorithm 1 terminates with the ITER-stop.

(B): $|I_G|=+\infty$, where $I_G\subseteq I$ denotes the index set that Algorithm 2 terminates with the GARD-stop.

Now, we begin by establishing the conclusion for Case (A). For simplicity, let us denote

$$\begin{cases}(M^\nu,C^\nu,K^\nu,G^\nu,N^\nu): \text{the sequence generated by Algorithm 2},\\ (M_*^\nu,C_*^\nu,K_*^\nu,G_*^\nu,N_*^\nu): \text{the sequence of optimal solutions for } (\mathscr{P}^\nu),\\ (M_*,C_*,K_*,G_*,N_*)\quad: \text{the optimal solution of } (\mathscr{P}).\end{cases}$$

It follows from the above definitions that

$$\begin{aligned}0&\le\Upsilon^\nu(M^\nu,C^\nu,K^\nu,G^\nu,N^\nu)\le\Upsilon^\nu(M_*^\nu,C_*^\nu,K_*^\nu,G_*^\nu,N_*^\nu)+\varepsilon^\nu\\&\le\Upsilon^\nu(M_*,C_*,K_*,G_*,N_*)+\varepsilon^\nu,\end{aligned}$$

which implies that

$$\begin{aligned}&\frac{1}{2}\|(M^\nu,C^\nu,K^\nu,G^\nu,N^\nu)-(M_0,C_0,K_0,G_0,N_0)\|^2\\&\le\frac{1}{2}\|(M_*,C_*,K_*,G_*,N_*)-(M_0,C_0,K_0,G_0,N_0)\|^2+\frac{\rho_\nu}{2}\sum_{i=1}^{2}\|\Gamma_i^\nu\|_F^2+\varepsilon^\nu/\rho_\nu \qquad (4.16)\\&\le\frac{1}{2}\|(M_*,C_*,K_*,G_*,N_*)-(M_0,C_0,K_0,G_0,N_0)\|^2+\frac{\rho_\nu}{2}\sum_{i=1}^{2}\|\Gamma_i^\nu\|_F^2+\rho_\nu\zeta. \qquad (4.17)\end{aligned}$$

Taking the limit of both sides of (4.17) in the index set I_I, we have

$$\begin{aligned}
&\frac{1}{2}\|(\bar{M},\bar{C},\bar{K},\bar{G},\bar{N})-(M_0,C_0,K_0,G_0,N_0)\|^2\\
&\le \frac{1}{2}\|(M_*,C_*,K_*,G_*,N_*)-(M_0,C_0,K_0,G_0,N_0)\|^2\\
&\quad + \lim_{\nu\in I_I}\frac{\rho_\nu}{2}\sum_{i=1}^{2}\|\Gamma_i^\nu\|_F^2+\lim_{\nu\in I_I}\rho_\nu\zeta\\
&= \frac{1}{2}\|(M_*,C_*,K_*,G_*,N_*)-(M_0,C_0,K_0,G_0,N_0)\|^2,
\end{aligned}$$

where the inequality follows from Lemma 4.1 and the last equation employs the fact that $\lim_{\nu\in I_I}\rho_\nu=0$. Hence, the conclusion holds for Case (A).

Next, we turn to prove the remaining part. Now, Algorithm 2 terminates with the GARD-stop. From Step 5 of Algorithm 2, we have

$$\begin{aligned}
&\|\rho_\nu(C^\nu-C_0)+\left[\mathscr{D}(\mathscr{T}(S,H_1^\nu))+\mathscr{B}(SH_2^\nu)\right]\|_F &&\le \tau^\nu,\\
&\|\rho_\nu(G^\nu-G_0)+\left[\mathscr{D}(\mathscr{T}^a(S,H_1^\nu))+\mathscr{B}^a(SH_2^\nu)\right]\|_F &&\le \tau^\nu,\\
&\|\rho_\nu(N^\nu-N_0)+\left[\mathscr{D}(\mathscr{T}^a(I_k,H_1^\nu))+\mathscr{B}^a(H_2^\nu)\right]\|_F &&\le \tau^\nu.
\end{aligned}$$

Due to the homogeneous property of the operators $\mathscr{D}$, $\mathscr{B}$, $\mathscr{B}^a$, $\mathscr{T}$, $\mathscr{T}^a$ and $H_i^\nu=-\rho_\nu\Gamma_i^{\nu+1}$, $(i=1,2)$, the above relations reduce to

$$\begin{aligned}
\|(C^\nu-C_0)-\left[\mathscr{D}(\mathscr{T}(S,\Gamma_1^{\nu+1}))+\mathscr{B}(S\Gamma_2^{\nu+1})\right]\|_F &\le \tau^\nu/\rho_\nu,\\
\|(G^\nu-G_0)-\left[\mathscr{D}(\mathscr{T}^a(S,\Gamma_1^{\nu+1}))+\mathscr{B}^a(S\Gamma_2^{\nu+1})\right]\|_F &\le \tau^\nu/\rho_\nu,\\
\|(N^\nu-N_0)-\left[\mathscr{D}(\mathscr{T}^a(I_k,\Gamma_1^{\nu+1}))+\mathscr{B}^a(\Gamma_2^{\nu+1})\right]\|_F &\le \tau^\nu/\rho_\nu. \qquad (4.18)
\end{aligned}$$

From Lemma 4.1, the sequence of $\{(\Gamma_1^\nu,\Gamma_2^\nu)\}$ is bounded. Hence, there exists a pair $(\bar{\Gamma}_1,\bar{\Gamma}_2)$ such that

$$\lim_{\nu\in I_G}(\Gamma_1^\nu,\Gamma_2^\nu)=(\bar{\Gamma}_1,\bar{\Gamma}_2).$$

The above inequalities and (4.18) tell us that

$$\begin{aligned}
\bar{C}-C_0&=\mathscr{D}(\mathscr{T}(S,\bar{\Gamma}_1))+\mathscr{B}(S\bar{\Gamma}_1),\\
\bar{G}-G_0&=\mathscr{D}(\mathscr{T}^a(S,\bar{\Gamma}_1))+\mathscr{B}^a(S\bar{\Gamma}_1),\\
\bar{N}-N_0&=\mathscr{D}(\mathscr{T}^a(I_k,\bar{\Gamma}_1))+\mathscr{B}^a(\bar{\Gamma}_1).
\end{aligned} \qquad (4.19)$$

Using the definition of GARD-stop again, the following relations hold at (M^ν,V^ν):

$$\begin{aligned}
&\|\Pi_{\mathscr{S}_+^n}\left(\rho_\nu M_0+L_F M^\nu-\left[\mathscr{D}(\mathscr{T}(S^2,H_1^\nu))+\mathscr{B}(S^2H_2^\nu)\right]\right)-(\rho_\nu+L_F)M^\nu\|_F\le\tau^\nu,\\
&\|\Pi_{\mathscr{S}_+^n}\left(\rho_\nu K_0+L_F M^\nu-\left[\mathscr{D}(\mathscr{T}(I_k,H_1^\nu))+\mathscr{B}(H_2^\nu)\right]\right)-(\rho_\nu+L_F)K^\nu\|_F\le\tau^\nu.
\end{aligned} \qquad (4.20)$$

Taking the limit of both sides of (4.20) and noting that $\rho_\nu \searrow 0$ as $\nu \to +\infty$, we have

$$\lim_{\nu \in I_G} \|\Pi_{\mathscr{S}^n_+}(\hat{M}^\nu) - M^\nu\|_F = 0, \quad \lim_{\nu \in I_G} \|\Pi_{\mathscr{S}^n_+}(\hat{K}^\nu) - K^\nu\|_F = 0, \tag{4.21}$$

where

$$\begin{cases} \hat{M}^\nu = M^\nu - \dfrac{\mathscr{D}\left(\mathscr{T}\left(S^2, H_1^\nu\right)\right) + \mathscr{B}\left(S^2 H_2^\nu\right)}{\rho_\nu + L_F} + \dfrac{\rho_\nu(M_0 - M^\nu)}{\rho_\nu + L_F}, \\ \hat{K}^\nu = K^\nu - \dfrac{\mathscr{D}\left(\mathscr{T}\left(I_k, H_1^\nu\right)\right) + \mathscr{B}\left(H_2^\nu\right)}{\rho_\nu + L_F} + \dfrac{\rho_\nu(K_0 - K^\nu)}{\rho_\nu + L_F}. \end{cases}$$

By the property of $\Pi_{\mathscr{S}^n_+}(\cdot)$ and (4.21), we have

$$\left\langle M - M^\nu, \hat{M}^\nu - M^\nu \right\rangle \le 0, \ \forall M \in \mathscr{S}^n_+, \quad \left\langle K - K^\nu, \hat{K}^\nu - K^\nu \right\rangle \le 0, \ \forall K \in \mathscr{S}^n_+,$$

which are equivalent to

$$\left\langle M - M^\nu, \frac{\hat{M}^\nu}{\rho_\nu} - \frac{M^\nu}{\rho_\nu} \right\rangle \le 0, \ \forall M \in \mathscr{S}^n_+, \quad \left\langle K - K^\nu, \frac{\hat{K}^\nu}{\rho_\nu} - \frac{K^\nu}{\rho_\nu} \right\rangle \le 0, \ \forall K \in \mathscr{S}^n_+. \tag{4.22}$$

Similar to the above discussions, it follows from $H_i^\nu = -\rho_\nu \Gamma_i^{\nu+1}$ $(i = 1, 2)$ that

$$\begin{cases} \dfrac{\hat{M}^\nu}{\rho_\nu} = \dfrac{M^\nu}{\rho_\nu} + \dfrac{\mathscr{D}\left(\mathscr{T}\left(S^2, \Gamma_1^{\nu+1}\right)\right) + \mathscr{B}\left(S^2 \Gamma_2^{\nu+1}\right)}{\rho_\nu + L_F} + \dfrac{(M_0 - M^\nu)}{\rho_\nu + L_F}, \\ \dfrac{\hat{K}^\nu}{\rho_\nu} = \dfrac{K^\nu}{\rho_\nu} + \dfrac{\mathscr{D}\left(\mathscr{T}\left(I_k, \Gamma_1^{\nu+1}\right)\right) + \mathscr{B}\left(\Gamma_2^{\nu+1}\right)}{\rho_\nu + L_F} + \dfrac{(K_0 - K^\nu)}{\rho_\nu + L_F}. \end{cases}$$

Hence, the relations (4.22) yields that

$$\begin{aligned} &\left\langle M^\nu - M_0 - \left(\mathscr{D}\left(\mathscr{T}\left(S^2, \Gamma_1^{\nu+1}\right)\right) + \mathscr{B}\left(S^2 \Gamma_2^{\nu+1}\right)\right), M - M^\nu \right\rangle \ge 0, \quad \forall M \in \mathscr{S}^n_+, \\ &\left\langle K^\nu - K_0 - \left(\mathscr{D}\left(\mathscr{T}\left(I_k, \Gamma_1^{\nu+1}\right)\right) + \mathscr{B}\left(\Gamma_2^{\nu+1}\right)\right), K - K^\nu \right\rangle \ge 0, \quad \forall K \in \mathscr{S}^n_+, \end{aligned}$$

which imply that

$$\begin{aligned} &\left\langle \bar{M} - M_0 - \left(\mathscr{D}\left(\mathscr{T}\left(S^2, \bar{\Gamma}_1\right)\right) + \mathscr{B}\left(S^2 \bar{\Gamma}_2\right)\right), M - \bar{M} \right\rangle \ge 0, \quad \forall M \in \mathscr{S}^n_+, \\ &\left\langle \bar{K} - K_0 - \left(\mathscr{D}\left(\mathscr{T}\left(I_k, \bar{\Gamma}_1\right)\right) + \mathscr{B}\left(\bar{\Gamma}_2\right)\right), K - \bar{K} \right\rangle \ge 0, \quad \forall K \in \mathscr{S}^n_+. \end{aligned} \tag{4.23}$$

From (4.19), (4.23) and Lemma 2.4, $(\bar{M}, \bar{C}, \bar{K}, \bar{G}, \bar{N})$ is an optimal solution of the following problem

$$\begin{array}{ll} \min & \Phi(M, C, K, G, N) - \langle \bar{\Gamma}_1, \mathscr{H}_1(M, C, K, G, N)\rangle - \langle \bar{\Gamma}_2, \mathscr{H}_2(M, C, K, G, N)\rangle \\ \text{s.t.} & (M, C, K, G, N) \in \Omega, \end{array} \tag{4.24}$$

In addition, $\mathscr{H}_1(\bar{M}, \bar{C}, \bar{K}, \bar{G}, \bar{N}) = 0$, $\mathscr{H}_2(\bar{M}, \bar{C}, \bar{K}, \bar{G}, \bar{N}) = 0$, because $(\bar{M}, \bar{C}, \bar{K}, \bar{G}, \bar{N})$ is a feasible solution of $(\mathscr{P})$. The conclusion for Case (B) follows from (4.24). □

Because the objective function of $(\mathscr{P})$ is a strictly convex function and the matrix Λ is nonsingular, the sequence generated by our algorithm ALA converges to the unique solution of $(\mathscr{P})$.

Corollary 4.3 *Let $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ be the sequence generated by Algorithm 2 with the parameters $\rho_\nu, \varepsilon^\nu, \tau^\nu$ satisfying the conditions mentioned in Lemma 4.2. Let $(M_*, C_*, K_*, G_*, N_*)$ be the unique solution of $(\mathscr{P})$. Then, we have*

$$\lim_{\nu\to\infty} (M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) = (M_*, C_*, K_*, G_*, N_*).$$

Next, we present an upper bound on the iteration-complexity for Algorithm 2.

Lemma 4.4 *Let $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ be the sequence generated by Algorithm 2 with the parameters $\rho_\nu, \varepsilon^\nu, \tau^\nu$ satisfying the conditions mentioned in Lemma 4.2. Let Σ^ν denote the number of iterations in Algorithm 1 for obtaining $(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)$. Then, there exists a positive constant Σ such that*

$$\Sigma^\nu \le \frac{\Sigma}{\sqrt{\varepsilon^\nu}}, \quad \forall \nu \ge 1. \tag{4.25}$$

Proof It follows from the definition of the ITER-stop that $\Sigma^\nu \le s^\nu_{\max}$. From (3.13), we have

$$\Sigma^\nu \le \sqrt{\frac{2L_F\xi^\nu}{\varepsilon^\nu}},$$

where $\xi^\nu := d^1(M_0, C_0, K_0, G_0, N_0) + d^\nu(M_0, C_0, K_0, G_0, N_0) + \rho_\nu \sum_{i=1}^2 \|\Gamma_i^\nu\|_F^2$. Combining the definition of d^ν in (3.10) with $(M^{\nu-1}, C^{\nu-1}, K^{\nu-1}, G^{\nu-1}, N^{\nu-1}) \in \mathscr{S}_\nu$, we obtain

$$d^\nu(M_0, C_0, K_0, G_0, N_0) \le 5(\eta^\nu)^2, \quad \forall \nu \ge 1.$$

where $\mathscr{S}_\nu, \eta^\nu$ are defined as (3.5–3.6), respectively. From (4.13) and Lemma 4.1, the relations $\eta^\nu \le B_\eta$ and $\sum_{i=1}^2 \|\Gamma_i^\nu\|_F^2 \le B_\eta^2$ are satisfied at the ν-th iteration. The above observations imply that

$$\Sigma^\nu \le \frac{\sqrt{2L_F\left(d+(5+\rho_1)B_\eta^2\right)}}{\sqrt{\varepsilon^\nu}} = \frac{\Sigma}{\sqrt{\varepsilon^\nu}},$$

where $\Sigma := \sqrt{2L_F\left(d+(5+\rho_1)B_\eta^2\right)}$. Hence, the conclusion (4.25) holds for all $\nu \ge 1$. □

The following lemma characterizes the infeasibility of the sequence generated by Algorithm 2.

Lemma 4.5 *Let $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ be the sequence generated by Algorithm 2 with the parameters $\rho_\nu, \varepsilon^\nu, \tau^\nu$ satisfying the conditions mentioned in Lemma 4.2. Then, we have*

$$\|\mathscr{H}_i(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\|_F \le \sqrt{2}B_\eta\rho_\nu, \quad (i=1,2), \tag{4.26}$$

where B_η is defined as (4.13).

Proof From Lemma 4.4 and (4.13), we obtain $\sum_{i=1}^2 \|\Gamma_i^\nu\|_F^2 \le B_\eta^2$. Notice that

$$\begin{aligned}
&\sum_{i=1}^2 \|\mathscr{H}_i(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\|_F^2 \\
&\quad \le \sum_{i=1}^2 \left(\|\mathscr{H}_i(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) - \rho_\nu\Gamma_i^\nu\|_F^2 + \|\rho_\nu\Gamma_i^\nu\|_F^2\right) \\
&\quad = \rho_\nu^2\left(\|\Gamma_1^{\nu+1}\|_F^2 + \|\Gamma_2^{\nu+1}\|_F^2 + \|\Gamma_1^\nu\|_F^2 + \|\Gamma_2^\nu\|_F^2\right) \le 2\rho_\nu^2 B_\eta^2,
\end{aligned}$$

which implies that the conclusion (4.26) holds. □

From Lemma 4.2, the sequence $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ is bounded and there exists a bound B_a such that

$$\max_\nu \left\{\|(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\|, \|(M_*, C_*, K_*, G_*, N_*)\|\right\} \le B_a. \tag{4.27}$$

Similarly, it follows from Lemma 4.1 that the sequence $\{(\Gamma_1^{\nu+1}, \Gamma_2^{\nu+1})\}$ is also bounded and there exists a bound B_o such that

$$\|\mathscr{D}(\mathscr{T}(S^2, \Gamma_1^{\nu+1})) + \mathscr{B}(S^2\Gamma_2^{\nu+1})\|_F + \|\mathscr{D}(\mathscr{T}(I_k, \Gamma_1^{\nu+1})) + \mathscr{B}(\Gamma_2^{\nu+1})\|_F \le B_o. \tag{4.28}$$

In the following lemma, we establish the sub-optimality of the sequence generated by Algorithm 2.

Lemma 4.6 *Let $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ be the sequence generated by Algorithm 2 with the parameters $\rho_\nu, \varepsilon^\nu, \tau^\nu$ satisfying the conditions mentioned in Lemma 4.2 and there exists a positive constant κ such that $\tau^\nu := \kappa\varepsilon^\nu$. Then, we have*

$$\begin{aligned}
&\frac{1}{2}\|(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\quad \le \frac{1}{2}\|(M_*, C_*, K_*, G_*, N_*) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\quad + \frac{B_\eta^2}{2}\rho_\nu + \sqrt{\zeta}\max\{1, \kappa(2B_\eta + B_o + 10B_a)\}\sqrt{\varepsilon^\nu},
\end{aligned}$$

where B_η is defined as (4.13), ζ is defined in Lemma 4.2 and B_a, B_o are defined as (4.27–4.28).

Proof Consider the following two cases:

(A) Algorithm 1 terminates with the ITER-stop. It follows from the inequality (4.16) that

$$\begin{aligned}
&\frac{1}{2}\|(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\le \frac{1}{2}\|(M_*, C_*, K_*, G_*, N_*) - (M_0, C_0, K_0, G_0, N_0)\|^2 + \frac{\rho_\nu}{2}\sum_{i=1}^{2}\|\Gamma_i^\nu\|_F^2 + \frac{\varepsilon^\nu}{\rho_\nu} \\
&\le \frac{1}{2}\|(M_*, C_*, K_*, G_*, N_*) - (M_0, C_0, K_0, G_0, N_0)\|^2 + \frac{B_\eta^2}{2}\rho_\nu + \frac{\varepsilon^\nu}{\rho_\nu}. \qquad (4.29)
\end{aligned}$$

(B) Algorithm 1 terminates with the GRAD-stop. Due to the convexity of Υ^ν, we obtain

$$\begin{aligned}
&\Upsilon^\nu(M_*, C_*, K_*, G_*, N_*) - \Upsilon^\nu(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) \\
&\quad \ge \left\langle \nabla\Upsilon^\nu(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu), (M_*, C_*, K_*, G_*, N_*) - (M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\right\rangle.
\end{aligned}$$

From (3.8) and (3.15), we have

$$\begin{aligned}
\nabla_M\Upsilon^\nu(M, C, K, G, N) &= \rho_\nu(M - M_0) + \mathscr{D}(\mathscr{T}(S^2, H_1^\nu)) + \mathscr{B}(S^2 H_2^\nu), \\
\nabla_C\Upsilon^\nu(M, C, K, G, N) &= \rho_\nu(C - C_0) + \mathscr{D}(\mathscr{T}(S, H_1^\nu)) + \mathscr{B}(S H_2^\nu), \\
\nabla_K\Upsilon^\nu(M, C, K, G, N) &= \rho_\nu(K - K_0) + \mathscr{D}(\mathscr{T}(I_k, H_1^\nu)) + \mathscr{B}(H_2^\nu), \\
\nabla_G\Upsilon^\nu(M, C, K, G, N) &= \rho_\nu(G - G_0) + \mathscr{D}(\mathscr{T}^a(S, H_1^\nu)) + \mathscr{B}^a(S H_2^\nu), \\
\nabla_N\Upsilon^\nu(M, C, K, G, N) &= \rho_\nu(N - N_0) + \mathscr{D}(\mathscr{T}^a(I_k, H_1^\nu)) + \mathscr{B}^a(H_2^\nu).
\end{aligned} \qquad (4.30)$$

Let us denote

$$\begin{cases}
\hat{M}^\nu = M^\nu - \dfrac{\mathscr{D}\left(\mathscr{T}\left(S^2, H_1^\nu\right)\right) + \mathscr{B}\left(S^2 H_2^\nu\right)}{\rho_\nu + L_F} + \dfrac{\rho_\nu(M_0 - M^\nu)}{\rho_\nu + L_F}, \\[2ex]
\hat{K}^\nu = K^\nu - \dfrac{\mathscr{D}\left(\mathscr{T}\left(I_k, H_1^\nu\right)\right) + \mathscr{B}(H_2^\nu)}{\rho_\nu + L_F} + \dfrac{\rho_\nu(K_0 - K^\nu)}{\rho_\nu + L_F}.
\end{cases}$$

By the property of $\Pi_{\mathscr{S}_+^n}(\cdot)$ and (4.21), we have

$$\left\langle \hat{M}^\nu - \Pi_{\mathscr{S}_+^n}\left(\hat{M}^\nu\right), M - \Pi_{\mathscr{S}_+^n}\left(\hat{M}^\nu\right)\right\rangle \leq 0, \quad \forall M \in \mathscr{S}_+^n,$$

which implies that

$$\begin{aligned}&\left\langle \hat{M}^\nu - M^\nu, M - M^\nu\right\rangle + \left\langle M^\nu - \Pi_{\mathscr{S}_+^n}\left(\hat{M}^\nu\right), M - M^\nu\right\rangle \\ &\quad + \left\langle \hat{M}^\nu - M^\nu, M^\nu - \Pi_{\mathscr{S}_+^n}\left(\hat{M}^\nu\right)\right\rangle \leq 0, \quad \forall M \in \mathscr{S}_+^n.\end{aligned}$$

It follows from the above inequality, (4.20) and (4.30) that

$$\begin{aligned}&-\left\langle \rho_\nu(M^\nu - M_0) + \mathscr{D}\left(\mathscr{T}\left(S^2, H_1^\nu\right)\right) + \mathscr{B}\left(S^2 H_2^\nu\right), M_* - M^\nu\right\rangle \leq \tau^\nu \|M_* - M^\nu\|_F \\ &\quad + \frac{\tau^\nu}{\rho_\nu + L_F}\|\rho_\nu(M^\nu - M_0) + \mathscr{D}\left(\mathscr{T}\left(S^2, H_1^\nu\right)\right) + \mathscr{B}\left(S^2 H_2^\nu\right)\|_F.\end{aligned}$$

Similarly, the following inequality holds at K^ν:

$$\begin{aligned}&-\left\langle \rho_\nu(K^\nu - K_0) + \mathscr{D}\left(\mathscr{T}\left(I_k, H_1^\nu\right)\right) + \mathscr{B}\left(H_2^\nu\right), K_* - K^\nu\right\rangle \leq \tau^\nu \|K_* - K^\nu\|_F \\ &\quad + \frac{\tau^\nu}{\rho_\nu + L_F}\|\rho_\nu(K^\nu - K_0) + \mathscr{D}\left(\mathscr{T}\left(I_k, H_1^\nu\right)\right) + \mathscr{B}\left(H_2^\nu\right)\|_F.\end{aligned}$$

From the above relations, we obtain

$$\begin{aligned}&\Upsilon^\nu(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) \\ &\quad \leq \Upsilon^\nu(M_*, C_*, K_*, G_*, N_*) + \tau^\nu(\|C_* - C^\nu\|_F + \|G_* - G^\nu\|_F + \|N_* - N^\nu\|_F) \\ &\qquad + \tau^\nu\|M_* - M^\nu\|_F + \frac{\tau^\nu}{\rho_\nu + L_F}\|\rho_\nu(M^\nu - M_0) + \mathscr{D}\big(\mathscr{T}\big(S^2, H_1^\nu\big)\big) + \mathscr{B}\big(S^2 H_2^\nu\big)\|_F \\ &\qquad + \tau^\nu\|K_* - K^\nu\|_F + \frac{\tau^\nu}{\rho_\nu + L_F}\|\rho_\nu(K^\nu - K_0) + \mathscr{D}\big(\mathscr{T}\big(I_k, H_1^\nu\big)\big) + \mathscr{B}\big(H_2^\nu\big)\|_F \\ &\quad \leq \Upsilon^\nu(M_*, C_*, K_*, G_*, N_*) + 10B_a\tau^\nu + \frac{\tau^\nu\rho_\nu}{\rho_\nu + L_F}\big\|(M^\nu - M_0) - \big(\mathscr{D}\big(\mathscr{T}\big(S^2, \Gamma_1^{\nu+1}\big)\big) \\ &\qquad + \mathscr{B}\big(S^2\Gamma_2^{\nu+1}\big)\big)\big\|_F \\ &\qquad + \frac{\tau^\nu\rho_\nu}{\rho_\nu + L_F}\big\|(K^\nu - K_0) - \big(\mathscr{D}\big(\mathscr{T}\big(I_k, \Gamma_1^{\nu+1}\big)\big) + \mathscr{B}\big(\Gamma_2^{\nu+1}\big)\big)\big\|_F \\ &\quad \leq \Upsilon^\nu(M_*, C_*, K_*, G_*, N_*) + 10B_a\tau^\nu + \tau^\nu 2B_\eta \\ &\qquad + \tau^\nu\left(\big\|\mathscr{D}\big(\mathscr{T}\big(S^2, \Gamma_1^{\nu+1}\big)\big) + \mathscr{B}\big(S^2\Gamma_2^{\nu+1}\big)\big\|_F + \big\|\mathscr{D}\big(\mathscr{T}\big(I_k, \Gamma_1^{\nu+1}\big)\big) + \mathscr{B}\big(\Gamma_2^{\nu+1}\big)\big\|_F\right) \\ &\quad \leq \Upsilon^\nu(M_*, C_*, K_*, G_*, N_*) + (2B_\eta + B_o + 10B_a)\tau^\nu, \end{aligned} \tag{4.31}$$

where the second inequality uses the fact that $H_i^\nu = -\rho_\nu\Gamma_i^{\nu+1}$ $(i = 1, 2)$ and the definition of B_a. The last inequality uses the definitions of B_η and B_o. Dividing (4.31) by ρ_ν, we obtain

$$\begin{aligned}
&\frac{1}{2}\|(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\quad \le \frac{1}{2}\|(M_*, C_*, K_*, G_*, N_*) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\qquad + \frac{\rho_\nu}{2}(\|\Gamma_1^\nu\|_F^2 + \|\Gamma_2^\nu\|_F^2) + (2B_\eta + B_o + 10B_a)\frac{\tau^\nu}{\rho_\nu} \\
&\quad \le \frac{1}{2}\|(M_*, C_*, K_*, G_*, N_*) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\qquad + \frac{B_\eta^2}{2}\rho_\nu + (2B_\eta + B_o + 10B_a)\frac{\tau^\nu}{\rho_\nu}.
\end{aligned} \tag{4.32}$$

It follows from (4.29) and (4.32) that

$$\begin{aligned}
&\frac{1}{2}\|(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\quad \le \frac{1}{2}\|(M_*, C_*, K_*, G_*, N_*) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\qquad + \frac{B_\eta^2}{2}\rho_\nu + \max\left\{\frac{\varepsilon^\nu}{\rho_\nu}, (2B_\eta + B_o + 10B_a)\frac{\tau^\nu}{\rho_\nu}\right\} \\
&\quad \le \frac{1}{2}\|(M_*, C_*, K_*, G_*, N_*) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\qquad + \frac{B_\eta^2}{2}\rho_\nu + \sqrt{\zeta}\max\{1, \kappa(2B_\eta + B_o + 10B_a)\}\sqrt{\varepsilon^\nu},
\end{aligned}$$

where the last inequality follows from $\varepsilon^\nu/(\rho_\nu^2) \le \zeta$ for all $\nu \ge 1$ and $\tau^\nu := \kappa\varepsilon^\nu$. □

We conclude this section in the following theorem about the convergence rate of Algorithm 2.

Theorem 4.2 *Let* $\{(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)\}$ *be the sequence generated by Algorithm 2 with the parameters* $\rho_\nu, \varepsilon^\nu, \tau^\nu$ *satisfying the conditions mentioned in Lemma 4.2 and there exist a positive constant* κ *and* $\sigma \in (0, 1)$ *such that*

$$\tau^\nu := \kappa\varepsilon^\nu, \ \ \rho_{\nu+1} := \sigma\rho_\nu, \ \ \varepsilon^{\nu+1} := \sigma^2\varepsilon^\nu. \tag{4.33}$$

Then, for any $\varepsilon > 0$*, Algorithm 2 needs at most*

$$\left\lceil \log_{\frac{1}{\sigma}}\left(\frac{\max\left\{\frac{B_\eta^2}{2}\rho_1 + \sqrt{\zeta}\max\{1, \kappa(2B_\eta + B_o + 10B_a)\}\sqrt{\varepsilon^1}, \sqrt{2}B_\eta\rho_1\right\}}{\varepsilon}\right)\right\rceil + 1$$

outer iterations and at most

$$\left\lceil \frac{\Sigma}{\sigma(1-\sigma)\sqrt{\varepsilon^1}} \left(\frac{\frac{B_\eta^2}{2}\rho_1 + \sqrt{\zeta}\max\{1, \kappa(2B_\eta + B_o + 10B_a)\}\sqrt{\varepsilon^1}}{\varepsilon} \right) \right\rceil$$

APG calls to obtain an ε-feasible and ε-optimal solution of ($\mathscr{P}$), where B_η is defined as (4.13), ζ is defined in Lemma 4.2 and B_a, B_o are defined as (4.27–4.28).

Proof From Lemma 4.6, we have

$$\begin{aligned}
&\frac{1}{2}\|(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\quad \le \frac{1}{2}\|(M_*, C_*, K_*, G_*, N_*) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\qquad + \frac{B_\eta^2}{2}\rho_\nu + \sqrt{\zeta}\max\{1, \kappa(2B_\eta + B_o + 10B_a)\}\sqrt{\varepsilon^\nu} \\
&\quad \le \frac{1}{2}\|(M_*, C_*, K_*, G_*, N_*) - (M_0, C_0, K_0, G_0, N_0)\|^2 \\
&\qquad + \left(\frac{B_\eta^2}{2}\rho_1 + \sqrt{\zeta}\max\{1, \kappa(2B_\eta + B_o + 10B_a)\}\sqrt{\varepsilon^1} \right)\sigma^{\nu-1},
\end{aligned}$$

which implies that $(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)$ becomes an ε-optimal solution of ($\mathscr{P}^\nu$) for all

$$\nu > \log_{\frac{1}{\sigma}} \left(\frac{\frac{B_\eta^2}{2}\rho_1 + \sqrt{\zeta}\max\{1, \kappa(2B_\eta + B_o + 10B_a)\}\sqrt{\varepsilon^1}}{\varepsilon} \right) + 1.$$

In addition, it follows from Lemma 4.5 that $(M^\nu, C^\nu, K^\nu, G^\nu, N^\nu)$ becomes an ε-feasible solution of ($\mathscr{P}$) for all

$$\nu > \log_{\frac{1}{\sigma}} \left(\frac{\sqrt{2}B_\eta\rho_1}{\varepsilon} \right) + 1.$$

Hence, Algorithm 2 needs at most

$$\left\lceil \log_{\frac{1}{\sigma}} \left(\frac{\max\left\{ \frac{B_\eta^2}{2}\rho_1 + \sqrt{\zeta}\max\{1, \kappa(2B_\eta + B_o + 10B_a)\}\sqrt{\varepsilon^1}, \sqrt{2}B_\eta\rho_1 \right\}}{\varepsilon} \right) \right\rceil + 1$$

outer iterations to obtain an ε-feasible and ε-optimal solution of ($\mathscr{P}$).

Now, we turn to the second part of the theorem. Let $\mathcal{N}_{outer}(\varepsilon;\kappa,\sigma)$ and $\mathcal{N}^t_{apg}(\varepsilon;\kappa,\sigma)$ denote the number of outer iterations and the total APG calls, respectively. Then, we have

$$\begin{aligned}\mathcal{N}^t_{apg}(\varepsilon;\kappa,\sigma) &\leq \sum_{\nu=1}^{\mathcal{N}_{outer}(\varepsilon;\kappa,\sigma)} \frac{\Sigma}{\sqrt{\varepsilon^\nu}} \\ &\leq \frac{\Sigma}{\sigma(1-\sigma)\sqrt{\varepsilon^1}}\left(\frac{\frac{B_\eta^2}{2}\rho_1+\sqrt{\zeta}\max\{1,\kappa(2B_\eta+B_o+10B_a)\}\sqrt{\varepsilon^1}}{\varepsilon}\right),\end{aligned}$$

where the first inequality follows from Lemma 4.4 and the last one is obtained from Lemma 4.6. □

5 Implementation details and numerical results

In this section, we conduct numerical experiments to test the performance of our proposed algorithm ALA in Sect. 3 by comparing it with the inexact smoothing Newton algorithm [21] and the interior-point method [16] for three inverse damped gyroscopic eigenvalue problems under different values of n, k. All experiments are performed in MATLAB 8.3.0 (2014a) 64-bit on a desktop computer with an Intel (R) Core(TM) 2 of 3.3 GHz CPU and 4 GB RAM running Windows 7.

5.1 Line search procedure for the subproblems in Remark 3.1

Now, we describe a line search procedure to implement Step 5 of Algorithm 2. From Remark 3.1, the subproblems $(\mathscr{P}^\nu_M)$ and $(\mathscr{P}^\nu_K)$ can be reformulated as follows:

$$\begin{cases}\min\ \lambda_1\|X-X_1\|_F^2+\lambda_2\|X-X_2\|_F^2 \\ \text{s.t.}\ \ \|X-X_1\|_F\leq\omega, \\ \qquad X\in\mathscr{S}^n_+,\end{cases} \tag{5.1}$$

where $X_1\in\mathscr{S}^n_+$, $X_2\in\mathscr{S}^n$ are two given different matrices and $\lambda_1,\lambda_2,\omega$ are positive scalars. The KKT conditions of (5.1) are given by

$$\begin{cases}0\in\lambda_1(X-X_1)+\lambda_2(X-X_2)+\mu(X-X_1)+\mathcal{N}_{\mathscr{S}^n_+}(X), \\ 0\leq\mu\perp(\|X-X_1\|_F^2-\omega^2)\leq 0.\end{cases} \tag{5.2}$$

Let us define the mappings $\hat{X}(\mu):[0,+\infty)\to\mathscr{S}^n$ and $g(\mu):[0,+\infty)\to\mathscr{R}$ as follows:

$$\hat{X}(\mu):=X_1+\frac{\lambda_2}{\lambda_1+\lambda_2+\mu}(X_2-X_1),\quad g(\mu)=\|\Pi_{\mathscr{S}^n_+}(\hat{X}(\mu))-X_1\|_F^2-\omega^2. \tag{5.3}$$

Hence, the relations (5.2) is equivalent to the following one-dimensional nonlinear complementarity problem:

$$0 \leq \mu \perp g(\mu) \leq 0. \tag{5.4}$$

The goal of the line search procedure for solving (5.4) can be described as follows:

Line search problem Given three positive scalars $\lambda_1, \lambda_2, \omega$ and the mapping $\hat{X}(\mu)$ as (5.3), the problem is to find $\mu^* \in [0, +\infty)$ satisfying the condition (5.4).

Notice that the following conclusions hold:

$$\mu^* = \begin{cases} 0 & \text{if } g(0) \leq 0, \\ \mu_0 & \text{if } g(0) > 0 \text{ and } g(\mu_0) = 0, \\ \mu_1 & \text{if } g(0) > 0 \text{ and } g(\mu_0) < 0, \end{cases}$$

where μ_0 is defined as follows:

$$\mu_0 := \frac{\lambda_2 \|X_2 - X_1\|_F}{\omega} - (\lambda_1 + \lambda_2)$$

and μ_1 is obtained by the bisection approach on $(0, \mu_0)$. From the property of $\Pi_{\mathscr{S}^n_+}(\cdot)$, we obtain

$$g(\mu) \leq \|\hat{X}(\mu) - X_1\|_F^2 - \omega^2 = \left(\frac{\lambda_2}{\lambda_1 + \lambda_2 + \mu}\right)^2 \|X_2 - X_1\|_F^2 - \omega^2,$$

which implies that $g(\mu_0) \leq 0$ and $\mu_0 > 0$ when $g(0) > 0$. Similar to the above discussion, line search procedures for solving another subproblems $(\mathscr{P}_C^\nu)$, $(\mathscr{P}_G^\nu)$ and $(\mathscr{P}_N^\nu)$ in Remark 3.1 can also be devised.

5.2 Implementation details

We describe the implementation details of Algorithm 2 in this subsection. In order to measure the infeasibility of $(\mathscr{P})$, let us denote

$$Res := \|M^\nu X \Lambda^2 + (C^\nu + G^\nu) X \Lambda + (K^\nu + N^\nu) X\|_F,$$

where the given partially measured eigenpair $(X, \Lambda) \in \mathscr{R}^{n\times k} \times \mathscr{R}^{k\times k}$ are generated by using MATLAB built-in functions '`rand`' and '`randn`', respectively. The total number of complex-valued eigenvalues is chosen to be around $k/2$. Similar to [21], we generate the measured eigenpair (X, Λ) repeatedly until X is full rank and Λ is nonsingular. In this paper, the ALA-stop for Algorithm 2 is defined as follows:

$$\max\{R_M^\nu, R_C^\nu, R_K^\nu, R_G^\nu, R_N^\nu\} \leq \mathtt{Tol},$$

where the notations $R_M^\nu, R_C^\nu, R_K^\nu, R_G^\nu, R_N^\nu$ are defined by

$$R_M^\nu := \|M^\nu - M^{\nu-1}\|_F,\ R_C^\nu := \|C^\nu - C^{\nu-1}\|_F,\ R_K^\nu := \|K^\nu - K^{\nu-1}\|_F,$$
$$R_G^\nu := \|G^\nu - G^{\nu-1}\|_F,\ R_N^\nu := \|N^\nu - N^{\nu-1}\|_F.$$

The pre-specified accuracy tolerance is denoted by `Tol` and set by default as $\texttt{Tol} = \log(n) \times 10^{-11}$. We set the initial penalty parameter ρ^1 as follows:

$$\rho^1 := \|M^0\|_2 + \|C^0\|_2 + \|N^0\|_2 + \|G^0\|_2 + \|N^0\|_2,$$

where $\|\cdot\|_2$ denotes the operator norm of a given matrix. The initial error bound ε^1 is given by

$$\varepsilon^1 := 0.5\rho^1 \| \left(M^0, C^0, K^0, G^0, N^0\right) - (M_0, C_0, K_0, G_0, N_0) \|^2$$

and the parameter $\sigma = 0.05$, where $(M^0, C^0, K^0, G^0, N^0)$ is the initial point of Algorithm 2, which is equal to $(\hat{M}, \hat{C}, \hat{K}, \hat{G}, \hat{N})$ defined in the Example 5.1 below. Let us define

$$m_F^{\nu,s} := \|TM_1^{\nu,s}\|_F,\ c_F^{\nu,s} := \|TC_1^{\nu,s}\|_F,\ k_F^{\nu,s} := \|TK_1^{\nu,s}\|_F,\ g_F^{\nu,s} := \|TG_1^{\nu,s}\|_F,$$
$$n_F^{\nu,s} := \|TN_1^{\nu,s}\|_F.$$

The parameter τ^ν is updated as follows:

$$\tau^\nu := \begin{cases} 0.999\left(m_F^{1,s} + c_F^{1,s} + k_F^{1,s} + g_F^{1,s} + n_F^{1,s}\right) & \text{if } \nu = 1, \\ \min\left\{m_F^{\nu,s} + 0.999c_F^{\nu,s} + k_F^{\nu,s} + 0.999g_F^{\nu,s} + 0.999n_F^{\nu,s}, \sigma\tau^{\nu-1}\right\} & \text{otherwise.} \end{cases}$$

In this paper, our APG subroutine for solving the subproblems in Remark 3.1 is similar to [29], whereas a line search procedure as above is added in our MATLAB codes.

5.3 Numerical examples

5.3.1 Synthetic sets

In this subsection, we report the performance of ALA for two synthetic datasets as in [21]. In all the tables below, the labels "it.", "CPU-time" and "Res" denote the number of iterations, the total computing time in seconds and the residual at the final iterate, respectively.

Example 5.1 Let $\hat{M}$, $\hat{C}$ and $\hat{K}$ be given as follows:

$$\hat{M} = \begin{bmatrix} R^{-T}R^{-1} & 0 \\ 0 & I \end{bmatrix},\ \hat{C} = \begin{bmatrix} -R^{-T}(\Lambda + \Lambda^T)R^{-1} & 0 \\ 0 & 0 \end{bmatrix},\ \hat{K} = \begin{bmatrix} R^{-T}\Lambda^T\Lambda R^{-1} & 0 \\ 0 & I \end{bmatrix},$$

$\hat{G}$ and $\hat{N}$ are two $n \times n$ null matrices. We define $(M_0, C_0, K_0, G_0, N_0)$ as follows:

$$M_0 := \hat{M} + \tau R_M, \quad C_0 := \hat{C} + \tau R_C, \quad K_0 := \hat{K} + \tau R_K,$$
$$G_0 := \hat{G} + \tau R_G, \quad N_0 := \hat{N} + \tau R_C,$$

where $R_M \in \mathscr{S}^n$, $R_C \in \mathscr{S}^n$, $R_K \in \mathscr{S}^n$ are three matrices with random entries uniformly distributed within $[-1.0, 1.0]$. R_G and R_N are two $n \times n$ skew-symmetric matrices with random entries between -1.0 and 1.0. The perturbed parameter is denoted by τ, here we set $\tau = 0.01, 0.1, 1.0$.

Remark 5.1 It follows from the definition of $(\hat{M}, \hat{C}, \hat{K}, \hat{G}, \hat{N})$ that

$$\mathscr{H}_1\left(\hat{M}, \hat{C}, \hat{K}, \hat{G}, \hat{N}\right) = 0, \quad \mathscr{H}_2\left(\hat{M}, \hat{C}, \hat{K}, \hat{G}, \hat{N}\right) = 0.$$

Example 5.2 The matrices $\hat{M}$, $\hat{K}$ are two random $n \times n$ correlation matrices generated by the MATLAB's gallery (`'randcorr'`,n). The matrix $\hat{C} \in \mathscr{S}^n$ with entries $\hat{C}_{ij} \in [-1.0, 1.0]$ and $\hat{C}_{ii} = 1.0$ for $i, j = 1, 2, \ldots, n$. The matrices $\hat{G}, \hat{N} \in \mathscr{K}^n$ are two random matrices with entries $\hat{G}_{ij}, \hat{N}_{ij} \in [-1.0, 1.0]$ and $\hat{G}_{ii} = 0$, $\hat{N}_{ii} = 0$ for $i, j = 1, 2, \ldots, n$. We define $(M_0, C_0, K_0, G_0, N_0)$ as follows:

$$\begin{aligned} &M_0 := \hat{M} + \tau R_M, \quad C_0 := \hat{C} + \tau R_C, \quad K_0 := \hat{K} + \tau R_K, \\ &G_0 := \hat{G} + \tau R_G, \quad N_0 := \hat{N} + \tau R_C, \end{aligned}$$

where $R_M \in \mathscr{S}^n$, $R_C \in \mathscr{S}^n$, $R_K \in \mathscr{S}^n$ are three matrices with random entries uniformly distributed within $[-1.0, 1.0]$. R_G and R_N are two $n \times n$ skew-symmetric matrices with random entries between -1.0 and 1.0. The perturbed parameter is denoted by τ, here we set $\tau = 0.01, 0.1, 1.0$.

5.3.2 Comparison with the interior-point method and the inexact smoothing Newton method

Now, we compare our algorithm ALA with the interior-point method [16] and the inexact smoothing Newton algorithm [21] for Examples 5.1 and 5.2. A state-of-art MATLAB package CVX/SDPT3 developed by Grant and Boyd [30] is used in our tests. The stopping criterions of the interior-point method and the inexact smoothing Newton algorithm are similar as the settings in [21].

Tables 1 and 2 show that the solvability of ALA and the interior-point method for Examples 5.1 and 5.2 for the instances with $n \leq 80$, but ALA is significantly

Table 1 Comparison of ALA and the interior-point method [16] for Example 5.1

τ	k	n	ALA			Interior-point method		
			It.	CPU-time	Res	It.	CPU-time	Res
1.0	5	40	11	0.5	1.37e−10	15	5.2	4.38e−07
		80	11	0.6	1.07e−10	16	47.8	5.09e−07
		120	11	1.3	4.86e−10	Out of memory		
		160	11	3.6	3.39e−10	–		
		200	11	5.1	8.86e−10	–		

Table 2 Comparison of ALA and the interior-point method [16] for Example 5.2

τ	k	n	ALA			Interior-point method		
			It.	CPU-time	Res	It.	CPU-time	Res
1.0	5	40	11	0.4	1.34e−10	14	4.8	8.79e−07
		80	11	1.0	9.85e−11	14	41.2	2.94e−09
		120	11	2.0	3.93e−10	Out of memory		
		160	11	2.5	6.48e−10	–		
		200	11	4.1	4.75e−10	–		

Table 3 Comparison of ALA and the inexact smoothing Newton method [21] for Example 5.1

τ	k	n	Newton			ALA		
			It.	CPU-time	Res	It.	CPU-time	Res
0.01	5	100	5	1.6	8.72e−11	10	1.2	3.14e−11
		200	5	5.8	1.34e−10	10	5.6	1.16e−10
		500	5	37.1	6.94e−10	10	29.8	3.01e−10
		1000	3	133.4	5.12e−08	9	104.6	2.19e−09
		1500	3	359.6	5.94e−08	9	346.3	5.60e−08
0.1	5	100	4	1.4	9.02e−10	10	1.4	2.98e−10
		200	5	5.8	4.68e−10	10	4.4	3.18e−10
		500	4	29.4	6.56e−09	10	22.8	5.16e−09
		1000	5	142.3	5.04e−09	10	130.5	5.01e−09
		1500	4	367.0	4.26e−08	10	364.2	1.45e−08
1.0	5	100	5	1.5	1.10e−09	11	2.2	2.49e-10
		200	5	5.6	1.14e−09	11	5.2	5.71e−10
		500	6	43.6	1.59e-09	11	35.7	1.51e−09
		1000	5	136.5	7.14e−09	10	122.8	6.31e−09
		1500	6	568.0	2.20e−09	11	547.0	1.19e−09

better than the interior-point method when $n \geq 120$. From the above observations, our method is more efficient in terms of CPU-time and can get solutions with lower residuals compared with the interior-point method.

Numerical results in Tables 3 and 4 indicate that our algorithm ALA is competitive with the inexact smoothing Newton method for all the instances under different values of τ, n, k. Compared with the inexact Newton method, our approach needs more iterations but costs less computing time and obtains finer solutions.

5.3.3 An inverse damped gyroscopic eigenvalue problem in structural engineering

In this subsection, we report the numerical experiments of ALA for an inverse damped gyroscopic eigenvalue problem in structural engineering, in which the mass matrix $\hat{M}$

Table 4 Comparison of ALA and the inexact smoothing Newton method [21] for Example 5.2

τ	k	n	Newton			ALA		
			It.	CPU-time	Res	It.	CPU-time	Res
0.01	5	100	6	2.0	3.04e−10	11	1.8	7.33e−11
		200	4	4.7	6.01e−10	11	3.2	2.36e−10
		500	4	28.5	4.81e−09	11	27.5	5.86e−10
		1000	6	203.8	6.80e−09	11	202.9	3.20e−09
		1500	6	405.3	4.41e−10	12	399.8	4.35e−10
0.1	5	100	4	1.3	8.77e−10	11	1.0	2.97e−11
		200	4	4.7	6.68e−10	11	3.8	4.68e−10
		500	4	28.3	2.35e−09	11	22.3	2.09e−09
		1000	5	195.6	4.77e−09	11	188.5	4.32e−09
		1500	5	438.7	2.16e−09	11	431.6	2.14e−09
1.0	5	100	5	1.6	6.63e−10	11	1.3	3.49e−10
		200	5	5.6	5.11e−10	10	4.8	4.79e−10
		500	5	34.6	3.89e−09	11	28.1	2.06e−09
		1000	5	142.2	4.04e−09	11	127.5	5.69e−09
		1500	6	446.4	2.83e−09	11	443.7	2.29e−09

and the stiffness matrix $\hat{K}$ come from the class BCSSTRUC1 in the Harwell-Boeing collection [31], the damping matrix $\hat{C}$ takes the form of $\hat{C} = \eta_0 \hat{M} + \eta_1 \hat{K}$, $(n = 66)$, where η_0 corresponds to the mass proportional Rayleigh damping parameter [32] and η_1 corresponds to the stiffness one.

Example 5.3 We set $\hat{M}$ and $\hat{K}$ to be BCSSTM02 and BCSSTK02, while $\hat{C} = 0.025\hat{M} + 0.025\hat{K}$, $\hat{G}$ and $\hat{N}$ are two null matrices. Similar to the above examples, we generate X repeatedly until X is full rank. The matrix Λ is a 12×12 diagonal matrix such that

$$\begin{aligned}\Lambda = \operatorname{diag}([&-40.5213, -40.4280, -40.4562, -40.4723, -40.4964, -40.4904,\\ &-40.3643, -40.2993, -40.2997, -40.3292, -40.3371, -40.3349]).\end{aligned}$$

The matrices M_0, C_0, K_0, G_0, N_0 are defined as follows:

$$\begin{aligned}&M_0 := \hat{M} + \tau R_M, \quad C_0 := \hat{C} + \tau R_C, \quad K_0 := \hat{K} + \tau R_K,\\ &G_0 := \hat{G} + \tau R_G, N_0 := \hat{N} + \tau R_C,\end{aligned}$$

where $R_M \in \mathscr{S}^n$, $R_C \in \mathscr{S}^n$, $R_K \in \mathscr{S}^n$ are three matrices with random entries uniformly distributed within $[-1.0, 1.0]$. R_G and R_N are two $n \times n$ skew-symmetric matrices with random entries between -1.0 and 1.0. The perturbed parameter is denoted by τ, here we set $\tau = 0.01, 0.05, 0.1, 0.5, 1.0$.

Table 5 shows that all three methods are comparable for Example 5.3. In addition, our algorithm ALA needs less iterations on average than the interior-point method

Table 5 Comparison of all three methods for Example 5.3

τ	Interior-point method			Newton			ALA		
	It.	CPU-time	Res	It.	CPU-time	Res	It.	CPU-time	Res
0.01	17	34.6	2.55e−06	9	2.7	3.14e−07	10	2.2	9.46e−08
0.05	18	36.5	1.19e−06	15	3.7	1.81e−07	10	2.7	1.45e−07
0.1	21	43.3	5.59e−06	14	3.1	4.43e−06	10	2.4	1.44e−07
0.5	19	39.2	9.92e−06	18	4.6	2.23e−06	9	2.9	2.21e−06
1.0	19	38.1	3.26e−06	19	4.5	3.03e−06	9	2.8	3.01e−06

and the inexact smoothing Newton method, especially for the instances with $\tau \geq 0.05$. Hence, for this inverse damped gyroscopic eigenvalue problem in structural engineering, our approach generally outperforms other solvers in terms of CPU-time and the corresponding residual.

6 Concluding remarks

In this paper, we consider an inverse damped gyroscopic eigenvalue problem. Under mild conditions, an augmented Lagrangian numerical method based on the APG strategy is proposed. We analyze the convergence of the proposed method and the corresponding iteration-complexity is established. Compared with other methods, our algorithm is easy to implement. Numerical experiments indicate that the proposed algorithm is an alternative approach for solving IDGEPs.

Acknowledgments This research was supported by the National Natural Science Foundation of China under Projects No. 91330206 and No. 91130007. The authors would like to thank the editor and the anonymous referees for their valuable comments for improving the presentation of this paper. We also thank Xi-jun Liang from China University of Petroleum and Ming Huang from Dalian University of Technology for helpful conversations and Xian-tao Xiao from Dalian University of Technology for contributions to code used in the numerical experiments.

7 Appendix 1

Let us denote

$$
\begin{aligned}
\mathscr{H}(M, C, K, G, N) &:= (\mathscr{H}_1(M, C, K, G, N), \mathscr{H}_2(M, C, K, G, N)), \ \Gamma^\nu := (\Gamma_1^\nu, \Gamma_2^\nu), \\
E &:= \Big([(S^2)^T \ 0^T] \ \ [S^T \ 0^T] \ \ [I_k^T \ 0^T] \ \ -[S^T \ 0^T] \ \ -[I_k^T \ 0^T] \Big), \\
\hat{E} &:= [S^2 \quad S \quad I_k \ \ -S \ \ -I_k].
\end{aligned}
$$

Hence, we obtain that $\mathscr{H}(M, C, K, G, N) = E[M^T, C^T, K^T, G^T, N^T]^T$. It follows from (3.7) that

$$
\nabla F^\nu(M, C, K, G, N) = \mathscr{H}^*(\mathscr{H}(M, C, K, G, N) - \rho_\nu \Gamma^\nu).
$$

Given two arbitrary pairs $(M_1, C_1, K_1, G_1, N_1)$, $(M_2, C_2, K_2, G_2, N_2) \in \Omega$, we have

$$
\begin{aligned}
&\|\nabla F^\nu(M_1, C_1, K_1, G_1, N_1) - \nabla F^\nu(M_2, C_2, K_2, G_2, N_2)\|_F^2 \\
&\quad = \|\mathscr{H}^*(\mathscr{H}(M_1, C_1, K_1, G_1, N_1) - \rho_\nu \Gamma^\nu) \\
&\qquad - \mathscr{H}^*(\mathscr{H}(M_2, C_2, K_2, G_2, N_2) - \rho_\nu \Gamma^\nu)\|_F^2 \\
&\quad = \|\mathscr{H}^*(\mathscr{H}(M_1 - M_2, C_1 - C_2, K_1 - K_2, G_1 - G_2, N_1 - N_2))\|_F^2 \\
&\quad = \|E^T E[M_1^T - M_2^T, C_1^T - C_2^T, K_1^T - K_2^T, G_1^T - G_2^T, N_1^T - N_2^T]^T\|_F^2 \\
&\quad \leq \|E^T E\|_F^2 \|(M_1, C_1, K_1, G_1, N_1) - (M_2, C_2, K_2, G_2, N_2)\|^2 \\
&\quad = \|\hat{E}^T \hat{E}\|_F^2 \|(M_1, C_1, K_1, G_1, N_1) - (M_2, C_2, K_2, G_2, N_2)\|^2
\end{aligned}
$$

$$= L_F^2\|(M_1, C_1, K_1, G_1, N_1) - (M_2, C_2, K_2, G_2, N_2)\|^2.$$

8 Appendix 2

It follows from the triangular inequality that

$$\begin{aligned}
&d^\nu(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)\\
&\quad= \|(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)-(M^{\nu-1}, C^{\nu-1}, K^{\nu-1}, G^{\nu-1}, N^{\nu-1})\|^2\\
&\quad\le \|(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)-(M_0, C_0, K_0, G_0, N_0)\|^2+d^\nu(M_0, C_0, K_0, G_0, N_0)\\
&\quad= \frac{2P^\nu(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)}{\rho_\nu} + d^\nu(M_0, C_0, K_0, G_0, N_0)\\
&\quad\le \frac{2\Upsilon^\nu(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)}{\rho_\nu} + d^\nu(M_0, C_0, K_0, G_0, N_0)\\
&\quad\le \frac{2\Upsilon^\nu(M^0, C^0, K^0, G^0, N^0)}{\rho_\nu} + d^\nu(M_0, C_0, K_0, G_0, N_0)\\
&\quad= \|(M^0, C^0, K^0, G^0, N^0) - (M_0, C_0, K_0, G_0, N_0)\|^2\\
&\qquad+ \rho_\nu \sum_{i=1}^{2} \|\Gamma_i^\nu\|_F^2 + d^\nu(M_0, C_0, K_0, G_0, N_0)\\
&\quad= d^1(M_0, C_0, K_0, G_0, N_0) + d^\nu(M_0, C_0, K_0, G_0, N_0) + \rho_\nu \sum_{i=1}^{2} \|\Gamma_i^\nu\|_F^2 =: \xi^\nu.
\end{aligned}$$

Hence, we have

$$s_{\max}^\nu = \sqrt{\frac{2L_F\xi^\nu}{\varepsilon^\nu}} \ge \sqrt{\frac{2L_F d^\nu(M_*^\nu, C_*^\nu, K_*^\nu, G_*^\nu, N_*^\nu)}{\varepsilon^\nu}}.$$

9 Appendix 3

Let $(\bar{M}^\nu, \bar{C}^\nu, \bar{K}^\nu, \bar{G}^\nu, \bar{N}^\nu) \in \Omega$ be an ε^ν-optimal solution of $(\mathscr{P}^\nu)$, that is,

$$\Upsilon^\nu(\bar{M}^\nu, \bar{C}^\nu, \bar{K}^\nu, \bar{G}^\nu, \bar{N}^\nu) \le \min\{\Upsilon^\nu(M, C, K, G, N) : (M, C, K, G, N) \in \Omega\} + \varepsilon^\nu.$$

From the definition of $\Upsilon^\nu(\cdot)$, we have

$$\begin{aligned}
&\Upsilon^\nu(M, C, K, G, N)\\
&\quad= \frac{\rho_\nu}{2}\|(M, C, K, G, N) - (M_0, C_0, K_0, G_0, N_0)\|^2\\
&\qquad+ \frac{1}{2}\sum_{i=1}^{2} \|\mathscr{H}_i(M, C, K, G, N) - \rho_\nu\Gamma_i^\nu\|_F^2
\end{aligned}$$

$$
\begin{aligned}
&\le \frac{\rho_\nu}{2}\left\|(M,C,K,G,N)-\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)\right\|^2\\
&\quad+\frac{\rho_\nu}{2}\left\|\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)-(M_0,C_0,K_0,G_0,N_0)\right\|^2\\
&\quad+\frac{1}{2}\sum_{i=1}^{2}\left\|\mathscr{H}_i\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)-\rho_\nu\Gamma_i^\nu\right\|_F^2\\
&\quad+\left\langle\nabla F^\nu\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right),(M,C,K,G,N)-\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)\right\rangle\\
&\quad+\frac{L_F}{2}\left\|(M,C,K,G,N)-\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)\right\|^2\\
&=\Upsilon^\nu\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)+\frac{\rho_\nu}{2}\left\|(M,C,K,G,N)-\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)\right\|^2\\
&\quad+\left\langle\nabla F^\nu\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right),(M,C,K,G,N)-\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)\right\rangle\\
&\quad+\frac{L_F}{2}\left\|(M,C,K,G,N)-\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)\right\|^2.
\end{aligned}
$$

Denote $(M_*^\nu,C_*^\nu,K_*^\nu,G_*^\nu,N_*^\nu)=\operatorname{argmin}\{\Upsilon^\nu(M,C,K,G,N):(M,C,K,G,N)\in\Omega\}$. It follows from the above relations and $(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu)\in\Omega$ that

$$
\begin{aligned}
&\Upsilon^\nu(M_*^\nu,C_*^\nu,K_*^\nu,G_*^\nu,N_*^\nu)\\
&\le\Upsilon^\nu\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)+\min\left\{\frac{\rho_\nu}{2}\left\|(M,C,K,G,N)-\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)\right\|^2\right.\\
&\quad+\left\langle\nabla F^\nu\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right),(M,C,K,G,N)-\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)\right\rangle\\
&\quad\left.+\frac{L_F}{2}\left\|(M,C,K,G,N)-\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)\right\|^2:(M,C,K,G,N)\in\Omega\right\}\\
&\le\Upsilon^\nu\left(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu\right)\\
&\quad+\min_{C\in\mathscr{S}^n}\left\{\frac{\rho_\nu}{2}\|C-\bar C^\nu\|_F^2+\langle\nabla_C F^\nu(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu),C-\bar C^\nu\rangle+\frac{L_F}{2}\|C-\bar C^\nu\|_F^2\right\}\\
&\quad+\min_{G\in\mathscr{K}^n}\left\{\frac{\rho_\nu}{2}\|G-\bar G^\nu\|_F^2+\langle\nabla_G F^\nu(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu),G-\bar G^\nu\rangle+\frac{L_F}{2}\|G-\bar G^\nu\|_F^2\right\}\\
&=\Upsilon^\nu(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu)-\frac{1}{2(\rho_\nu+L_F)}\left(\|\nabla_C\ F^\nu(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu)\|_F^2\right.\\
&\quad\left.-\|\nabla_G F^\nu(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu)\|_F^2\right)\\
&\le\Upsilon^\nu(M_*^\nu,C_*^\nu,K_*^\nu,G_*^\nu,N_*^\nu)+\varepsilon^\nu-\frac{1}{2(\rho_\nu+L_F)}\left(\|\nabla_C F^\nu(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu)\|_F^2\right.\\
&\quad\left.-\|\nabla_G F^\nu(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu)\|_F^2\right),
\end{aligned}
$$

which implies that

$$
\begin{aligned}
&\|\nabla_C F^\nu(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu)\|_F^2\\
&\quad+\|\nabla_G F^\nu(\bar M^\nu,\bar C^\nu,\bar K^\nu,\bar G^\nu,\bar N^\nu)\|_F^2\le 2(\rho_\nu+L_F)\varepsilon^\nu.
\end{aligned}
$$

References

1. Tisseur, F., Meerbergen, K.: The quadratic eigenvalue problem. SIAM Rev. **43**, 235–286 (2001)
2. Sheena, Z., Unger, A., Zalmanovich, A.: Theoretical stiffness matrix correction by static test results. Isr. J. Tech. **20**, 245–253 (1982)
3. Baruch, M.: Optimization procedure to correct stiffness and flexibility matrices using vibration data. AIAA J. **16**, 1208–1210 (1978)
4. Baruch, M., Bar-Itzack, I.Y.: Optimal weighted othogonalization of measured modes. AIAA J. **16**, 346–351 (1978)
5. Berman, A.: Comment on optimal weighted othogonalization of measured modes. AIAA J. **17**, 927–928 (1979)
6. Berman, A., Nagy, E.J.: Improvement of a large analytical model using test data. AIAA J. **21**, 1168–1173 (1983)
7. Caesar, B.: Update and identification of dynamic mathematical models. 4th IMAC, Los Angeles, CA, pp. 394–401 (1986)
8. Wei, F.S.: Structural dynamic model identification using vibration test data. 7th IMAC, Las Vegas, Nevada, pp. 562–567 (1989)
9. Wei, F.S.: Structural dynamic model improvement using vibration test data. AIAA J. **28**, 175–177 (1990)
10. Wei, F.S.: Mass and stiffness interaction effects in analytical model modification. AIAA J. **28**, 1686–1688 (1990)
11. Friswell, M., Inman, D., Pilkey, D.: The direct updating of damping and stiffness matrices. AIAA J. **36**, 491–493 (1998)
12. Lancaster, P.: Strongly stable gyroscopic systems. Electron. J. Linear Algebra **5**, 53–66 (2008)
13. Bai, Z., Chu, D., Sun, D.: A dual optimization approach to inverse quadratic eigenvalue problems with partial eigenstructure. SIAM J. Sci. Comput. **29**, 2531–2561 (2007)
14. Chu, D., Chu, M., Lin, W.: Quadratic model updating with symmetric, positive definiteness, and no spill-over. SIAM J. Matrix Anal. Appl. **31**, 546–564 (2009)
15. Chu, M.T., Buono, N.D.: Total decoupling of a general quadratic pencil, part i: theory. J. Sound Vib. **309**, 96–111 (2008)
16. Lin, M., Dong, B., Chu, M.T.: Semidefinite programming techniques for structured quadratic inverse eigenvalue problems. Numer. Algorithms **53**, 419–437 (2010)
17. Jia, Z.G., Wei, M.S.: A real-valued spectral decomposition of the undamped gyroscopic system with applications. SIAM J. Matrix Anal. Appl. **32**, 584–604 (2011)
18. Ammari, M.A., Tisseur, F.: Standard triples of structured matrix polynomials. Linear Algebra Appl. **437**, 817–834 (2012)
19. Lancaster, P.: Stability of linear gyroscopic systems: A review. Linear Algebra Appl. **439**, 686–706 (2013)
20. Chu, M.T., Kuo, Y.C., Lin, W.W.: On inverse quadratic eigenvalue problems with partially prescribed eigenstructure. SIAM J. Matrix Anal. Appl. **25**, 995–1020 (2004)
21. Xiao, X.T., Gu, J., Zhang, L.W.: Quadratic model updating with gyroscopic structure from partial eigendata. Optim. Eng. **14**, 431–455 (2013)
22. Tseng, P.: On accelerated proximal gradient methods for convex-concave optimization. http://pages.cs.wisc.edu/~brecht/cs726docs/Tseng.APG
23. Toh, K.C., Yun, S.W.: An accelerated proximal gradient algorithm for nuclear norm regularized least squares problems. Pac. J. Optim. **6**, 615–640 (2010)
24. Jiang, K.F., Sun, D.F., Toh, K.C.: An inexact accelerated proximal gradient method for large scale linearly constrained convex SDP. SIAM J. Optim. **22**, 1042–1064 (2012)
25. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**, 183–202 (2009)
26. Tseng, P.: Approximation accuracy, gradient methods, and error bound for structured convex optimization. Math. Program. **125**, 263–295 (2010)
27. Nesterov, Y.: Introductory Lectures on Convex Optimization: A Basic Course. Springer, Berlin (2004)
28. Rockafellar, R., Wets, R.: Variational Analysis. Springer, New York (1998)
29. MATLAB codes for APG. https://github.com/bodono/apg/blob/master/apg.m
30. Grant, M., Boyd, S.: CVX: Matlab software for disciplined convex programming, version 2.1 (2014). http://cvxr.com/cvx

31. The Harwell-Boeing Collection. http://math.nist.gov/MatrixMarket/data/Harwell-Boeing
32. Friswell, M., Mottershead, J.: Finite Element Model Updating in Structural Dynamics. Kluwer, Dordrecht (1995)