RESEARCH ARTICLE
doi: 10.2306/scienceasia1513-1874.2021.104

Iterative methods for the quadratic bilinear equation arising from a class of quadratic dynamic systems

Bo Yu, Ning Dong*, Qiong Tang

School of Science, Hunan University of Technology, Zhuzhou 412008 China

*Corresponding author, e-mail: dongning_158@sina.com

Received 25 Feb 2021
Accepted 3 Sep 2021

ABSTRACT: To solve the quadratic bilinear equation arising from the dynamical system, a fixed-point iterative method and Newton's method are considered in this paper. The iteration sequence generated by two methods, starting from the zero matrix is proved to be monotonically increasing and convergent to the minimal positive (semi-)definite solution. Besides, a double Newton step is given to accelerate the current Newton's iteration when the equation is near or in the semi-stable case. Numerical experiments demonstrate the effectiveness of the fixed-point iteration and Newton's method with the ADI preconditioning. In particular, the adapted double Newton step can efficiently decrease iterative steps of Newton's method when the equation is semi-stable.

KEYWORDS: quadratic bilinear system, fixed-point iteration, Newton's method, ADI preconditioning, double Newton step

MSC2010: 65F10 15A24

INTRODUCTION

Consider a class of quadratic bilinear equations with the Hadamard product (QBEH)

$$\mathcal{Q}(X)=AX+XA^{\mathrm{T}}+MXM^{\mathrm{T}}+(GXG^{\mathrm{T}})\circ(FXF^{\mathrm{T}})+D=0 \tag{1}$$

arising from the dynamical system

$$\begin{aligned}\dot{x}(t)&=Ax(t)+H(x(t)\otimes x(t))+Mx(t)u(t)+Bu(t),\\ y(t)&=Cx(t),\end{aligned} \tag{2}$$

where M, G, F, D in (1) are real matrices of order n and D is symmetric positive (semi-)definite, $x(t)\in\mathbb{R}^n$ in (2) represents the state vector at time t, $u(t)\in\mathbb{R}^m$ and $y(t)\in\mathbb{R}$ denote the input and the output functions, respectively, $A\in\mathbb{R}^{n\times n}$, $B\in\mathbb{R}^{n\times m}$ $(m<n)$ and $C\in\mathbb{R}^{1\times n}$ stand for the state matrix, input matrix and the output matrix, respectively, $H\in\mathbb{R}^{n\times n^2}$ is a sparse matrices associated with the quadratic functions $x(t)\otimes x(t)$ [1]. In some systems such as the transmission line circuit, comprising of resistors, capacitors, and diodes [1–4], the matrix H has some implicit structure and the quadratic item $H(x(t)\otimes x(t))$ could be represented via the Hadamard product $(Gx(t))\circ(Fx(t))$. Take the sparse matrix H in the system of transmission line circuit as an example [3]. The elements are of the following structure

$$\begin{aligned}H(\tfrac{n}{2}+i,(i-1)n+\tfrac{n}{2}+i)&=H(\tfrac{n}{2}+i,\tfrac{n^2}{2}+(i-1)n+\tfrac{n}{2}+i)\\ &=-80 \text{ for } i=2,\dots,\tfrac{n}{2},\\ H(\tfrac{n}{2}+i,in+\tfrac{n}{2}+i)&=H(\tfrac{n}{2}+i,\tfrac{n^2}{2}+in+\tfrac{n}{2}+i)\\ &=40 \text{ for } i=2,\dots,\tfrac{n}{2}-1,\\ H(\tfrac{n}{2}+i,(i-2)n+\tfrac{n}{2}+i)&=H(\tfrac{n}{2}+i,\tfrac{n^2}{2}+(i-2)n+\tfrac{n}{2}+i)\\ &=-80 \text{ for } i=3,\dots,\tfrac{n}{2}\\ H(\tfrac{n}{2}+1,\tfrac{n}{2}+1)&=H(\tfrac{n}{2}+1,\tfrac{n^2}{2}+\tfrac{n}{2}+1)=-40,\\ H(\tfrac{n}{2}+1,\tfrac{3n}{2}+1)&=H(\tfrac{n}{2}+1,\tfrac{n^2}{2}+\tfrac{3n}{2}+1)=-40,\\ H(\tfrac{n}{2}+2,\tfrac{n}{2}+2)&=H(\tfrac{n}{2}+2,\tfrac{n^2}{2}+\tfrac{n}{2}+2)=-40\end{aligned}$$

and other are zeros. Let $x(t)=(v_1,v_{12},\dots,v_{n-1,n},y_1,\dots,y_n)^{\mathrm{T}}\in\mathbb{R}^{2n}$, $G=I_{2n}$ and

$$F=\begin{bmatrix}0_n & 0_n\\ T & T\end{bmatrix}\in\mathbb{R}^{2n\times 2n}$$

with tri-diagonal matrix $T=\text{tridiag}(40,-80,40)$, $T(1,1)=-40$, $T(1,2)=-40$ and $T(2,1)=-40$, the quadratic item $H(x(t)\otimes x(t))$ could be represented as $(Gx(t))\circ(Fx(t))$. Then the system (2) further develops to the quadratic bilinear system with Hadamard product (QBSH) [5]

$$\begin{aligned}\dot{x}(t)&=Ax(t)+(Gx(t))\circ(Fx(t))+Mx(t)u(t)+Bu(t),\\ y(t)&=Cx(t).\end{aligned} \tag{3}$$

To describe the controllability and observability of the system effectively, the Gramian matrix of the system, i.e. the solution to the QBEH (1) with $D = BB^{\mathrm{T}}$ is required. If A is stable and there is a positive (semi-)definite matrix Z such that $\mathscr{Q}(Z) \geqslant 0$, the existence of the solution is then guaranteed via employing a fixed-point iteration [5]. However in most cases, the initial point selection X_0 satisfying $X_0 \geqslant Z$ and $\mathscr{Q}(X_0) \leqslant 0$ is normally not easy and might make the fixed-point iteration hard to realize. In this paper, a more convenient choice of the initial point, as well as the following contributions are provided.

(i) The fixed-point iteration is reconsidered with an initial zero matrix at fingertips, producing a monotonically increasing sequence that converges to the minimal positive (semi-)definite solution to the QBEH (1);

(ii) Newton's method with the initial zero matrix is proposed to solve the QBEH and it shares the analogously monotonic convergence with that of the fixed-point iteration. Especially, the alternated-direction-implicit (ADI) iteration [6, 7] is employed as the preconditioning to compute Newton's subproblem and the selection strategy of ADI parameters is given in detail;

(iii) When the QBEH (1) is semi-stable, the convergence rate of Newton's method degenerates to be linear and can be re-accelerated by implementing a double Newton step [8, 9].

Numerical experiments indicate that the fixed-point method and the devised Newton-ADI preconditioning method are effective solvers to calculate the minimal positive (semi-)definite solution to the QBEH (1). Meanwhile, the double Newton step is very efficient to accelerate the original Newton's method when the QBEH (1) is near to or in the semi-stable case.

Throughout this paper, it is written $A \geqslant B$ ($A > B$) for symmetric matrices A and B if $A - B$ is a symmetric positive semi-definite (definite) matrix. By $\mathbb{R}_{+}^{n\times n} = \{A \in \mathbb{R}^{n\times n} \,|\, A \geqslant 0\}$ we denote the closed convex cone of non-negative definite matrices. Symbols $\sigma(A)$, $\rho(A) = \max\{|\alpha| : \alpha \in \sigma(A)\}$ and $\lambda(A) = \max\{Re(\alpha) : \alpha \in \sigma(A)\}$ are the spectrum, the spectral radius and the spectral abscissa of the matrix A, respectively. Several definitions and lemmas are also required.

Definition 1 ([10]) The matrix A is called stable or semi-stable if its spectrum lies in the left half of the complex plane $\mathbb{C}_{<}$, or the left half of the complex plane plus the imaginary axis $\mathbb{C}_{\leqslant}$.

Definition 2 ([11]) A linear operator $\mathscr{L} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is called positive if it maps $\mathbb{R}_{+}^{n\times n}$ to $\mathbb{R}_{+}^{n\times n}$ and inverse positive if $\mathscr{L}^{-1}$ exists and is positive. When the operator $\alpha I - \mathscr{L}$ is inverse positive for sufficiently large $\alpha > 0$, the operator $\mathscr{L}$ is called resolvent positive.

Lemma 1 ([12]) *Let $\mathscr{L} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be a resolvent positive linear operator and $\mathscr{T} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be a positive linear operator. Then the adjoint operator of $\mathscr{L}$, denoted by $\widetilde{\mathscr{L}}$, is also resolvent positive and the following assertions hold:*

i. There exists a matrix $V \geqslant 0, V \neq 0$ such that $\mathscr{L}(V) = \lambda(\mathscr{L})V$.

ii. If $\mathscr{S}$ is a linear operator with $\mathscr{S} \geqslant \mathscr{L}$, then $\mathscr{S}$ is resolvent positive and $\lambda(\mathscr{S}) \geqslant \lambda(\mathscr{L})$.

iii. $aI - \mathscr{L}$ is inverse positive $\Leftrightarrow$ $a > \lambda(\mathscr{L})$ $\Leftrightarrow$ $\sigma(\mathscr{L} - aI) \subset \mathbb{C}_{<}^{n\times n}$.

iv. $\mathscr{L} + \mathscr{T}$ is stable if and only if there exists $X > 0$ satisfying $(\mathscr{L} + \mathscr{T})(X) < 0$.

Lemma 2 ([12]) *Let the operator $\mathscr{F}$ be Fréchet differentiable on an open neighbourhood of a convex subset of $\mathbb{R}^{n\times n}$. If for some $V \geqslant 0, V \neq 0$*

$$\langle V, \mathscr{F}(Y) - \mathscr{F}(X)\rangle = \langle V, \mathscr{F}_X'(Y - X)\rangle,$$

then $\langle V, \mathscr{F}_X'(\cdot) - \mathscr{F}_Y'(\cdot)\rangle = 0$, i.e. $\widetilde{\mathscr{F}}_X'(V) = \widetilde{\mathscr{F}}_Y'(V)$. Here $\langle\cdot\rangle$ represents the inner product, $\widetilde{\mathscr{F}}$ stands for the adjoint operator of $\mathscr{F}$ and $\mathscr{F}_X'(\cdot)$ is the Fréchet derivative at X.

Lemma 3 ([13]) *Let the matrix $A \in \mathbb{R}^{n\times n}$ be stable and $D \in \mathbb{R}^{n\times n}$ be symmetric. Then the Lyapunov equation*

$$AX + XA^{\mathrm{T}} + D = 0$$

has a unique symmetric solution X. Moreover, $X \geqslant 0$ if $D \geqslant 0$.

Lemma 4 ([14]) *Let $A, B \in \mathbb{R}^{n\times n}$ be symmetric matrices.*

1. If $A > 0$ and $B > 0$, then $A \circ B > 0$.

2. If $A \geqslant 0$ and $B \geqslant 0$, then $A \circ B \geqslant 0$. Moreover, $A \circ B > 0$ when A has no zero row.

Lemma 5 ([15]) *Let $A, B \in \mathbb{C}^{n\times n}$ be complex matrices and $\|\cdot\|$ be any unitarily invariant norm. Then*

$$\|A \circ B\|^2 \leqslant \|A^*A\|\|B^*B\|$$

with "$(\cdot)^*$" being the conjugate transpose of a matrix.

ITERATIVE METHODS FOR THE QBEH

Fixed-point iteration

When matrix A is stable and there exists a matrix $Z \geqslant 0$ such that $\mathscr{Q}(Z) \geqslant 0$, it has been shown that there is a solution (actually the maximal positive (semi-)definite solution) to the QBEH (1) [5]. The way to prove is based on a fixed-point iteration

$$\mathscr{L}(X_{k+1}) = -(GX_kG^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}}) - MX_kM^{\mathrm{T}} - D, \quad k \geqslant 0 \quad (4)$$

with an initial matrix $X_0 \geqslant Z$ and $\mathscr{Q}(X_0) \leqslant 0$, where $\mathscr{L}$ is a linear operator $\mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ given by

$$\mathscr{L}(X) = AX + XA^{\mathrm{T}}$$

and has been shown resolvent positive [12]. Generally, finding such an initial matrix is not easy even if the solution exists. In this section, we reconsider the fixed-point iteration (4) but with a different zero initial point, i.e. $X_0 = 0$. The produced iteration sequence $\{X_k\}$ will be demonstrated to be monotonically increasing and converging to the minimal positive (semi-)definite solution to the QBEH (1).

Theorem 1 *Let A be a stable matrix. Let $X^* \geqslant 0$ be the solution to the QBEH* (1)*. The fixed-point iteration* (4) *with $X_0 = 0$ produces a matrix sequence $\{X_k\}$ such that for $k \geqslant 0$, $X_k \leqslant X_{k+1}$, $X_k \leqslant X^*$, $\mathscr{Q}(X_k) \geqslant 0$. Then the QBEH* (1) *has a minimal positive (semi-)definite solution $\hat{X}$.*

Proof: Starting with $X_0 = 0$, the fixed-point iteration obviously yields $X_1 \geqslant 0 = X_0$ and $\mathscr{Q}(X_0) = D \geqslant 0$. Now suppose that

$$X_i \leqslant X_{i+1}, \quad X_i \leqslant X^*, \quad \mathscr{Q}(X_i) \geqslant 0 \qquad (5)$$

holds for $i = k$, we shall show that it is valid for $i = k+1$. Firstly, by the iteration format (4) and Lemma 4, we have

$$\mathscr{L}(X^* - X_{k+1}) = -(GX^*G^{\mathrm{T}})\circ(F(X^* - X_k)F^{\mathrm{T}}) \\ -(G(X^* - X_k)G^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}}) - M(X^* - X_k)M^{\mathrm{T}} \leqslant 0$$

and

$$\mathscr{L}(X_{k+2} - X_{k+1}) = -(GX_kG^{\mathrm{T}})\circ(F(X_{k+1} - X_k)F^{\mathrm{T}}) \\ -(G(X_{k+1} - X_k)G^{\mathrm{T}})\circ(FX_{k+1}F^{\mathrm{T}}) - M(X_{k+1} - X_k)M^{\mathrm{T}} \leqslant 0,$$

respectively. Then $X_{k+1} \leqslant X^*$ and $X_{k+1} \leqslant X_{k+2}$ hold by using Lemma 1. This goes together with

$$\begin{aligned}\mathscr{Q}(X_{k+1}) &= AX_{k+1} + X_{k+1}A^{\mathrm{T}} + (GX_{k+1}G^{\mathrm{T}})\circ(FX_{k+1}F^{\mathrm{T}}) \\ &\quad + MX_{k+1}M^{\mathrm{T}} + D \\ &= -A(X_{k+2} - X_{k+1}) - (X_{k+2} - X_{k+1})A^{\mathrm{T}} \\ &= -\mathscr{L}(X_{k+2} - X_{k+1}) \geqslant 0\end{aligned}$$

show that the induction assumption (5) holds for $i = k+1$. Then the sequence $\{X_k\}$ is well defined and has a limit $\lim_{k\to\infty} X_k = \hat{X}$ such that $\hat{X} \leqslant X^*$. Moreover, $\hat{X}$ is the minimal positive (semi-)definite solution. □

Remark 1 Conditions of the stable A and $Q(Z) \geqslant 0$ are required to prove the existence of the positive (semi-)definite solution (actually the maximal positive (semi-)definite solution) to the QBEH [5]. With an easier initial matrix (i.e. $X_0 = 0$), Theorem 1 shows the existence of the minimal positive (semi-)definite solution to the QBEH. Although two extreme solutions might be different, they are both of main concerned in practices.

Newton's method

Let the first order and the second order Fréchet derivative of $\mathscr{Q}(\cdot)$ at X be

$$\mathscr{Q}'_X(\Delta) = A\Delta + \Delta A^{\mathrm{T}} + M\Delta M^{\mathrm{T}} \\ + (G\Delta G^{\mathrm{T}})\circ(FXF^{\mathrm{T}}) + (GXG^{\mathrm{T}})\circ(F\Delta F^{\mathrm{T}}) \quad (6)$$

and

$$\mathscr{Q}''_X(\Delta_1, \Delta_2) = (G\Delta_1G^{\mathrm{T}})\circ(F\Delta_2F^{\mathrm{T}}) \\ + (G\Delta_2G^{\mathrm{T}})\circ(F\Delta_1F^{\mathrm{T}}),$$

respectively. Given the initial matrix X_0, for $k = 0, 1, \ldots$, Newton's method

$$X_{k+1} = X_k - \mathscr{Q}'_{X_k}(\mathscr{Q}(X_k))$$

admits the following iteration

$$\begin{aligned}AX_{k+1} + X_{k+1}A^{\mathrm{T}} + MX_{k+1}M^{\mathrm{T}} + (GX_{k+1}G^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}}) \\ + (GX_kG^{\mathrm{T}})\circ(FX_{k+1}F^{\mathrm{T}}) \\ = (GX_kG^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}}) - D. \quad (7)\end{aligned}$$

Theorem 2 *Let $X^* \geqslant 0$ be the solution to the QBEH* (1)*. Newton's iteration iteration* (7) *with $X_0 = 0$ produces a matrix sequence $\{X_k\}$ such that for $k \geqslant 0$,*

1. *$X_k \leqslant X_{k+1}$, $X_k \leqslant X^*$, $\mathscr{Q}(X_k) \geqslant 0$, $\mathscr{Q}'_{X_k}$ is stable;*
2. *$\lim_{k\to\infty} X_k = \hat{X}$ is a positive (semi-)definite solution to the QBEH* (1)*. Especially, $\hat{X}$ is the minimal positive (semi-)definite solution.*

Proof: The theorem is also proved by the induction applied to

$$X_i \leqslant X_{i+1}, X_i \leqslant X^*, \mathscr{Q}(X_i) \geqslant 0, \mathscr{Q}'_{X_i} \text{ is stable}, \ i \geqslant 0. \quad (8)$$

Starting with $X_0 = 0$, one has $\mathscr{Q}(X_0) \geqslant 0$ and

$$\mathscr{Q}'_{X_0} = \mathscr{L} + \mathscr{T}$$

with $\mathscr{L}(\cdot) = A(\cdot) + (\cdot)A^{\mathrm{T}}$ and $\mathscr{T}(\cdot) = M(\cdot)M^{\mathrm{T}}$. It then follows the iteration (7) and Lemma 1 that $\mathscr{Q}'_{X_0}$ is stable. Moreover, $X_1 = X_0 - (\mathscr{Q}'_{X_0})^{-1}\mathscr{Q}(X_0) \geqslant X_0$, so the induction (8) holds for $i = 0$.

Assume that (8) is true for $i = k$. We next show the case for $i = k+1$. Firstly, it follows from the iteration format (7) that

$$\begin{aligned}
\mathscr{Q}'_{X_k}(X^* - X_{k+1}) &= A(X^* - X_{k+1}) + (X^* - X_{k+1})A^{\mathrm{T}} \\
&\quad + (G(X^* - X_{k+1})G^{\mathrm{T}}) \circ (FX_kF^{\mathrm{T}}) \\
&\quad + (GX_kG^{\mathrm{T}}) \circ (F(X^* - X_{k+1})F^{\mathrm{T}}) + M(X^* - X_{k+1})M^{\mathrm{T}} \\
&= D - (GX_kG^{\mathrm{T}}) \circ (FX_kF^{\mathrm{T}}) + AX^* + X^*A^{\mathrm{T}} \\
&\quad + (GX^*G^{\mathrm{T}}) \circ (FX_kF^{\mathrm{T}}) + (GX_kG^{\mathrm{T}}) \circ (FX^*F^{\mathrm{T}}) + MX^*M^{\mathrm{T}} \\
&= -(G(X^* - X_k)G^{\mathrm{T}}) \circ (F(X^* - X_k)F^{\mathrm{T}}) \leqslant 0.
\end{aligned} \tag{9}$$

As $\mathscr{Q}'_{X_k}$ is stable, $-(\mathscr{Q}'_{X_k})^{-1}$ is positive. So $X_{k+1} \leqslant X^*$.

Next we show that $\mathscr{Q}'_{X_{k+1}}$ is stable. Let

$$\begin{aligned}
&\mathscr{L} = A(\cdot) + (\cdot)A^{\mathrm{T}}, \\
&\mathscr{M}_{X_{k+1}} = (G(\cdot)G^{\mathrm{T}}) \circ (FX_{k+1}F^{\mathrm{T}}) + (GX_{k+1}G^{\mathrm{T}}) \circ (F(\cdot)F^{\mathrm{T}}).
\end{aligned}$$

Then it follows from Lemma 1(ii) that $\mathscr{Q}'_{X_{k+1}} = \mathscr{L} + \mathscr{M}_{X_{k+1}}$ is resolvent positive and thus the adjoint operator $\widetilde{\mathscr{Q}}'_{X_{k+1}}$ is also resolvent positive. So by Lemma 1, there exists $V \geqslant 0$ such that

$$\widetilde{\mathscr{Q}}'_{X_{k+1}} V = \lambda V, \quad (\lambda \geqslant 0). \tag{10}$$

If $\mathscr{Q}'_{X_{k+1}}$ is assumed to be unstable, then

$$\langle V, \mathscr{Q}'_{X_{k+1}}(X^* - X_{k+1})\rangle = \langle \lambda V, X^* - X_{k+1}\rangle \geqslant 0.$$

On the other hand,

$$\begin{aligned}
\mathscr{Q}'_{X_{k+1}}(X^* - X_{k+1}) &= A(X^* - X_{k+1}) + (X^* - X_{k+1})A^{\mathrm{T}} \\
&\quad + (G(X^* - X_{k+1})G^{\mathrm{T}}) \circ (FX_{k+1}F^{\mathrm{T}}) \\
&\quad + (GX_{k+1}G^{\mathrm{T}}) \circ (F(X^* - X_{k+1})F^{\mathrm{T}}) + M(X^* - X_{k+1})M^{\mathrm{T}} \\
&= \mathscr{Q}(X^*) - \mathscr{Q}(X_{k+1}) - (G(X_{k+1} - X^*)G^{\mathrm{T}}) \circ (F(X_{k+1} - X^*)F^{\mathrm{T}}) \\
&\leqslant -(G(X_{k+1} - X^*)G^{\mathrm{T}}) \circ (F(X_{k+1} - X^*)F^{\mathrm{T}}) \\
&\quad - (G(X_{k+1} - X_k)G^{\mathrm{T}}) \circ (F(X_{k+1} - X_k)F^{\mathrm{T}}) \\
&\leqslant -(G(X_{k+1} - X_k)G^{\mathrm{T}}) \circ (F(X_{k+1} - X_k)F^{\mathrm{T}}) \leqslant 0
\end{aligned}$$

indicates that $\langle V, \mathscr{Q}'_{X_{k+1}}(X^* - X_{k+1})\rangle \leqslant 0$. So one has $\langle V, \mathscr{Q}'_{X_{k+1}}(X^* - X_{k+1})\rangle = 0$, implying

$$\langle V, (G(X_{k+1} - X_k)G^{\mathrm{T}}) \circ (F(X_{k+1} - X_k)F^{\mathrm{T}})\rangle = 0.$$

Note that

$$\begin{aligned}
\mathscr{Q}'_{X_k}(X_{k+1} - X_k) &= \mathscr{Q}(X_{k+1}) - \mathscr{Q}(X_k) \\
&\quad - (G(X_{k+1} - X_k)G^{\mathrm{T}}) \circ (F(X_{k+1} - X_k)F^{\mathrm{T}}).
\end{aligned}$$

Then

$$\langle V, \mathscr{Q}(X_{k+1}) - \mathscr{Q}(X_k)\rangle = \langle V, \mathscr{Q}'_{X_k}(X_{k+1} - X_k)\rangle.$$

It follows from Lemma 2 that

$$\begin{aligned}
\widetilde{\mathscr{Q}}'_{X_k}(V) &= \widetilde{\mathscr{Q}}'_{X_{k+1}}(V) + \widetilde{\mathscr{Q}}'_{X_k - X_{k+1}}(V) \\
&= \widetilde{\mathscr{Q}}'_{X_{k+1}}(V) = \lambda V \geqslant 0,
\end{aligned}$$

contradicting the stability of $\mathscr{Q}'_{X_k}$. So $\mathscr{Q}'_{X_{k+1}}$ is stable.

Now, by Newton's iteration (7), one has

$$\begin{aligned}
\mathscr{Q}'_{X_{k+1}}(X_{k+2} - X_{k+1}) &= A(X_{k+2} - X_{k+1}) + (X_{k+2} - X_{k+1})A^{\mathrm{T}} \\
&\quad + (G(X_{k+2} - X_{k+1})G^{\mathrm{T}}) \circ (FX_{k+1}F^{\mathrm{T}}) \\
&\quad + (GX_{k+1}G^{\mathrm{T}}) \circ (F(X_{k+2} - X_{k+1})F^{\mathrm{T}}) + M(X_{k+2} - X_{k+1})M^{\mathrm{T}} \\
&= -AX_{k+1} - X_{k+1}A^{\mathrm{T}} - (GX_{k+1}G^{\mathrm{T}}) \circ (FX_{k+1}F^{\mathrm{T}}) \\
&\quad - MX_{k+1}M^{\mathrm{T}} - D \\
&= (GX_{k+1}G^{\mathrm{T}}) \circ (FX_kF^{\mathrm{T}}) + (GX_kG^{\mathrm{T}}) \circ (FX_{k+1}F^{\mathrm{T}}) \\
&\quad - (GX_{k+1}G^{\mathrm{T}}) \circ (FX_{k+1}F^{\mathrm{T}}) - (GX_kG^{\mathrm{T}}) \circ (FX_kF^{\mathrm{T}}) \\
&= -(G(X_{k+1} - X_k)G^{\mathrm{T}}) \circ (F(X_{k+1} - X_k)F^{\mathrm{T}}) \leqslant 0.
\end{aligned}$$

As $\mathscr{Q}'_{X_{k+1}}$ is stable, $-(\mathscr{Q}'_{X_{k+1}})^{-1}$ is positive. So $X_{k+2} \geqslant X_{k+1}$.

Finally, Newton's iteration $\mathscr{Q}(X_{k+1}) = -\mathscr{Q}'_{X_{k+1}}(X_{k+2} - X_{k+1}) \geqslant 0$ shows that the induction holds for $i = k+1$. Therefore, the sequence $\{X_k\}$ is well defined, monotonically increasing and bounded by X^*. So $\lim_{k\to\infty} X_k = \hat{X}$. Moreover, $\hat{X} \leqslant X^*$ shows that $\hat{X}$ is the minimal positive (semi-)definite solution to the QBEH (1). □

The following theorem further indicates the quadratic convergence of Newton's iteration.

Theorem 3 *Let the sequence $\{X_k\}$ be produced by Newton's iteration (7) and $\hat{X}$ be the minimal positive definite solution to the QBEH. If $\mathscr{Q}'_{\hat{X}}$ is stable, then there is a constant θ such that*

$$\|X_{k+1} - \hat{X}\| \leqslant \theta \|X_k - \hat{X}\|^2,$$

where $\|\cdot\|$ is any unitarily invariant norm.

Proof: It follows from the Newton's iteration (7) that

$$\begin{aligned}\mathcal{Q}'_{\hat{X}}(\hat{X}-X_{k+1}) &= A(\hat{X}-X_{k+1})+(\hat{X}-X_{k+1})A^{\mathrm{T}}\\ &\quad +(G(\hat{X}-X_{k+1})G^{\mathrm{T}})\circ(F\hat{X}F^{\mathrm{T}})\\ &\quad +(G\hat{X}G^{\mathrm{T}})\circ(F(\hat{X}-X_{k+1})F^{\mathrm{T}})+M(\hat{X}-X_{k+1})M^{\mathrm{T}}\\ &= (GX_{k+1}G^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}})+(GX_kG^{\mathrm{T}})\circ(FX_{k+1}F^{\mathrm{T}})\\ &\quad -(GX_kG^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}})+(G\hat{X}G^{\mathrm{T}})\circ(F\hat{X}F^{\mathrm{T}})\\ &\quad -(G\hat{X}G^{\mathrm{T}})\circ(FX_{k+1}F^{\mathrm{T}})-(GX_{k+1}G^{\mathrm{T}})\circ(F\hat{X}F^{\mathrm{T}})\\ &= (G(\hat{X}-X_{k+1})G^{\mathrm{T}})\circ(F(\hat{X}-X_{k+1})F^{\mathrm{T}})\\ &\quad -(G(X_{k+1}-X_k)G^{\mathrm{T}})\circ(F(X_{k+1}-X_k)F^{\mathrm{T}}).\end{aligned}$$

Then one has

$$\begin{aligned}&\|\hat{X}-X_{k+1}\| \leqslant\\ &\quad \|(\mathcal{Q}'_{\hat{X}})^{-1}\|\cdot\|(G(\hat{X}-X_{k+1})G^{\mathrm{T}})\circ(F(\hat{X}-X_{k+1})F^{\mathrm{T}})\|.\end{aligned}$$

Note that $\|(\mathcal{Q}'_{\hat{X}})^{-1}\|$ is bounded above, it then follows from Lemma 5 that there exists a constant $\theta > 0$ such that

$$\|\hat{X}-X_{k+1}\| \leqslant \theta\|X_{k+1}-X_k\|^2 \leqslant \theta\|\hat{X}-X_k\|^2,$$

where the last inequality comes from the monotonic convergence of Newton's sequence $\{X_k\}$ for $k \geqslant 0$. □

Theorem 3 indicates that the convergence rate of Newton's method (7) is quadratic when $\mathcal{Q}'_{\hat{X}}$ is stable. If it is semi-stable, the convergence will degenerate to be linear. So the acceleration of the iteration (7) can be further considered. Before that, we concentrate on the calculation of Newton's subproblem.

ADI preconditioning for the linear subproblem

To efficiently implement Newton's method, a linear matrix equation with Hadamard product

$$AX+XA^{\mathrm{T}}+MXM^{\mathrm{T}}+(GXG^{\mathrm{T}})\circ(F_X)+(G_X)\circ(FXF^{\mathrm{T}}) = E \quad (11)$$

requires to be solved, here F_X, G_X and E are available symmetric positive definite matrices at the current step. Directly solving (11) is not easy but again a fixed-point form

$$AX+XA^{\mathrm{T}} = \hat{D} \quad (12)$$

is feasible with the current available $\hat{D} = E - MXM^{\mathrm{T}} - (GXG^{\mathrm{T}})\circ(F_X) - (G_X)\circ(FXF^{\mathrm{T}})$. This is a standard Lyapunov equation and might be further accelerated by proper preconditioning. Here a cyclic Smith or ADI preconditioning as in [16, 17] is employed. Specifically, by incorporating the ADI parameters $p_l > 0$ for $l \geqslant 0$ and rewriting the QBEH as

$$\begin{aligned}&AX+XA^{\mathrm{T}}\\ &\quad = \frac{1}{2p_l}\Big((A-p_lI)X(A-p_lI)^{\mathrm{T}}-(A+p_lI)X(A+p_lI)^{\mathrm{T}}\Big),\end{aligned}$$

the iteration (12) will implement in the following way

$$\begin{aligned}&Y_{k,0} = X_k,\\ &Y_{k,l} = \tilde{A}_{p_l}Y_{k,l-1}\tilde{A}^{\mathrm{T}}_{p_l}+2p_lA_{p_l}\hat{D}_kA^{\mathrm{T}}_{p_l},\ 1\leqslant l\leqslant L,\\ &X_{k+1} = Y_{k,L},\end{aligned} \quad (13)$$

where

$$\begin{aligned}&\tilde{A}_{p_l} = (A-p_lI)^{-1}(A+p_lI), \quad A_{p_l} = (A-p_lI)^{-1},\\ &\hat{D}_k = E-MX_kM^{\mathrm{T}}-(GX_kG^{\mathrm{T}})\circ(F_X)-(G_X)\circ(FX_kF^{\mathrm{T}}).\end{aligned}$$

Generally, given the prescribed accuracy, the ADI iteration number L will be available by Wachspress's method [6, 7] before starting the iterations. Then the iteration scheme (13) becomes

$$\begin{aligned}X_{k+1} &= Y_{k,L}\\ &= \tilde{A}_{p_L}Y_{k,L-1}\tilde{A}^{\mathrm{T}}_{p_L}+2p_LA_{p_L}\hat{D}_kA^{\mathrm{T}}_{p_L}\\ &= \tilde{A}_{p_L}(\tilde{A}_{p_{L-1}}Y_{k,L-2}\tilde{A}^{\mathrm{T}}_{p_{L-1}}+2p_{L-1}A_{p_{L-1}}\hat{D}_kA^{\mathrm{T}}_{p_{L-1}})\tilde{A}^{\mathrm{T}}_{p_L}\\ &\quad +2p_LA_{p_L}\hat{D}_kA^{\mathrm{T}}_{p_L}\\ &\quad \vdots\\ &= \Big(\prod_{l=1}^{L}\tilde{A}_{p_l}\Big)X_k\Big(\prod_{l=1}^{L}\tilde{A}_{p_l}\Big)^{\mathrm{T}}\\ &\quad +\sum_{l=1}^{L}2p_l\Big(\Big(\prod_{i=l+1}^{L}\tilde{A}_{p_i}A_{p_l}\Big)\hat{D}_k\Big(\prod_{i=l+1}^{L}\tilde{A}_{p_i}A_{p_l}\Big)^{\mathrm{T}}\Big).\end{aligned} \quad (14)$$

The fixed-point of the above iteration is the solution to the equation (11) and the iteration process essentially transforms the equation (11) into a preconditioned equation

$$\begin{aligned}&\sum_{l=1}^{L}2p_l\Big(\prod_{i=l+1}^{L}\tilde{A}_{p_i}A_{p_l}\Big)(S(X))\Big(\prod_{i=l+1}^{L}\tilde{A}_{p_i}A_{p_l}\Big)^{\mathrm{T}}\\ &\qquad = -\sum_{l=1}^{L}2p_l\Big(\Big(\prod_{i=l+1}^{L}\tilde{A}_{p_i}A_{p_l}\Big)E\Big(\prod_{i=l+1}^{L}\tilde{A}_{p_i}A_{p_l}\Big)^{\mathrm{T}}\Big).\end{aligned}$$

with $S(X) = AX+XA^{\mathrm{T}}+MXM^{\mathrm{T}}+(GXG^{\mathrm{T}})\circ(F_X)+(G_X)\circ(FXF^{\mathrm{T}})$ and parameters p_l being independent of the iteration number k.

It is known that the choice of parameters p_l in $\prod_{l=1}^{L}\tilde{A}_{p_l}$ determines the convergence rate of the ADI preconditioning and it is solved by Wachspress via the Jacobian elliptic function.

Let $\epsilon >$ be the tolerance and $a, b > 0$ be constants such that $\sigma(A) \subset [a, b]$. The minimal ADI iteration number L and the ADI parameters can be determined by the following min-max problem

$$\min_{p_1,\dots,p_L}\max_{\gamma\in[a,b]}\prod_{l=1}^{L}\left|\frac{\gamma-p_l}{\gamma+p_1}\right| \leqslant \epsilon.$$

Define the complete elliptic integral of the first kind $u = F(\phi, m) = \int_0^\phi \frac{1}{\sqrt{1-m\sin^2\theta}}\,\mathrm{d}\theta$ and the corresponding elliptic function $\mathrm{d}n(u) = \sqrt{1-m\sin^2\theta}$. Then

$$L = \left[\sqrt{\frac{F(\pi/2,\sqrt{1-(a/b)^2})}{2\pi F(\pi/2,a/b)}\ln\frac{4}{\epsilon}}\right]$$

with $[x]$ denoting the smallest integer toward the floor and

$$p_l = b \times \mathrm{d}n\Big(\frac{2l-1}{2L}F\Big(\frac{\pi}{2},\sqrt{1-(\frac{a}{b})^2}\Big)\Big).$$

Remark 2 As the computation of L and p_l in ADI preconditioning is independent of the fixed-point iteration and Newton's iteration, it can be determined before the iteration process when A and the desired accuracy are available. Then the fixed-point iteration of the form (14) is applied to the preconditioned subproblem.

SEMI-STABLE CASE

The quadratic convergence of Newton's method generally degenerates to be linear when the QBEH (1) is in the semi-stable case. In this case, a double step as in [8, 9] can be employed to improve the performance of Newton's iteration (7).

Let $\mathscr{N}$ be the null space of the Fréchet operator $\mathscr{Q}'_{\hat{X}}$ and $\mathscr{M}$ be the corresponding complement space in $\mathbb{R}^{n\times n}$. Let $P_{\mathscr{N}}$ and $P_{\mathscr{M}}$ be the orthogonal projections onto $\mathscr{N}$ and $\mathscr{M}$, respectively. Normally, for sufficiently large k, if the error of Newton's sequence $X_k - \hat{X}$ tends to lie in the space $\mathscr{M}$ rather than the space $\mathscr{N}$, i.e.,

$$\|P_{\mathscr{M}}(X_k - \hat{X})\| > c\|P_{\mathscr{N}}(X_k - \hat{X})\|$$

for some constant c. Then Newton's iteration (7) might be quadratically convergent [8]. On the contrary, if the error is dominated by the null space $\mathscr{N}$, the convergent rate is not quadratic as the following theorem states.

Theorem 4 *Suppose that Newton's method (7) produces the sequence $\{X_k\}$, converging to the solution $\hat{X}$. If $\mathscr{Q}'_{\hat{X}}$ is semi-stable and $X_k - \hat{X} \in \mathscr{N}$, then*

1. $\mathscr{Q}'_{X_k}(X_k - \hat{X}) = 2\mathscr{Q}(X_k)$.
2. $X_{k+1} - \hat{X} = \frac{1}{2}(X_k - \hat{X})$.

Proof: Note that $X_k - \hat{X} \in \mathscr{N}$ implies

$$\begin{aligned}&A(X_k - \hat{X}) + (X_k - \hat{X})A^{\mathrm{T}} + M(X_k - \hat{X})M^{\mathrm{T}}\\&\quad = (G\hat{X}G^{\mathrm{T}})\circ(F(\hat{X}-X_k)F^{\mathrm{T}}) + (G(\hat{X}-X_k)G^{\mathrm{T}})\circ(F\hat{X}F^{\mathrm{T}}).\end{aligned}$$

This together with (6) yields

$$\begin{aligned}\mathscr{Q}'_{X_k}(X_k - \hat{X}) &= A(X_k - \hat{X}) + (X_k - \hat{X})A^{\mathrm{T}} + M(X_k - \hat{X})M^{\mathrm{T}}\\
&\quad +(GX_kG^{\mathrm{T}})\circ(F(X_k-\hat{X})F^{\mathrm{T}})+(G(X_k-\hat{X})G^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}})\\
&= \mathscr{Q}(X_k) - A\hat{X} - \hat{X}A^{\mathrm{T}} - M\hat{X}M^{\mathrm{T}} - (GX_kG^{\mathrm{T}})\circ(F\hat{X}F^{\mathrm{T}})\\
&\quad - (G\hat{X}G^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}}) + (GX_kG^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}})\\
&= \mathscr{Q}(X_k) + (G\hat{X}G^{\mathrm{T}})\circ(F(\hat{X}-X_k)F^{\mathrm{T}})\\
&\quad + (G(\hat{X}-X_k)G^{\mathrm{T}})\circ(F\hat{X}F^{\mathrm{T}})\\
&\quad - (G\hat{X}G^{\mathrm{T}})\circ(F\hat{X}F^{\mathrm{T}}) + (GX_kG^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}})\\
&= \mathscr{Q}(X_k) + A(X_k - \hat{X}) + (\hat{X} - X_k)A^{\mathrm{T}} + M(X_k - \hat{X})X^{\mathrm{T}}\\
&\quad - (G\hat{X}G^{\mathrm{T}})\circ(F\hat{X}F^{\mathrm{T}}) + (GX_kG^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}})\\
&= 2\mathscr{Q}(X_k) + \mathscr{Q}(\hat{X})\\
&= 2\mathscr{Q}(X_k).\end{aligned}$$

So assertion 1 holds true. By Newton's iteration (7), one has

$$\begin{aligned}\mathscr{Q}'_{X_k}(X_{k+1}-\hat{X}) &= A(X_{k+1}-\hat{X})+(X_{k+1}-\hat{X})A^{\mathrm{T}}+M(X_{k+1}-\hat{X})M^{\mathrm{T}}\\
&\quad +(GX_kG^{\mathrm{T}})\circ(F(X_{k+1}-\hat{X})F^{\mathrm{T}})+(G(X_{k+1}-\hat{X})G^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}})\\
&= (GX_kG^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}}) - D - A\hat{X} - \hat{X}A^{\mathrm{T}} - M\hat{X}M^{\mathrm{T}}\\
&\quad - (GX_kG^{\mathrm{T}})\circ(F\hat{X}F^{\mathrm{T}}) - (G\hat{X}G^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}})\\
&= \mathscr{Q}'_{X_k}(X_k - \hat{X}) - \mathscr{Q}(X_k).\end{aligned}$$

Then assertion 2 holds via $\mathscr{Q}'_{X_k}(X_k - \hat{X}) = 2\mathscr{Q}(X_k)$. □

When $\mathscr{Q}'_{\hat{X}}$ is semi-stable, Theorem 4 indicates that the iteration error will be dominated by the null space $\mathscr{N}$ of $\mathscr{Q}'_{\hat{X}}$, with the linear convergence of the constant $1/2$. Moreover, assertion 1 in Theorem 4 provides a simple accelerated strategy when X_k is close to solution $\hat{X}$, i.e., the double Newton step

$$\begin{aligned}&AX_{k+1}+X_{k+1}A^{\mathrm{T}}+MX_{k+1}M^{\mathrm{T}}+(GX_kG^{\mathrm{T}})\circ(FX_{k+1}F^{\mathrm{T}})\\&\quad + (GX_{k+1}G^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}})\\&\quad = -AX_k - X_kA^{\mathrm{T}} - MX_kM^{\mathrm{T}} - 2D. \qquad (15)\end{aligned}$$

In our practical implementations, Newton's method will be switched to the double Newton step when the iterative sequence is close to the solution to the semi-stable QBEH.

NUMERICAL EXPERIMENTS

In this section, we demonstrate the effectiveness of the proposed iterative methods for solving the QBEH (1). The fixed-point iteration scheme (4) ("FIP"), Newton's method (7) with ADI preconditioning ("NEW") and the Hybrid Newton's method with a double Newton step ("DN-NEW") were coded by MATLAB 2014 and all examples were operated on a laptop with an Intel i3-3240 3.4GHz processor and 8GB RAM. The terminated condition in each algorithm is that the relative residual satisfied ReQX < tol or the iteration exceeds 100, where

$$\mathrm{ReQX}=\frac{\|AX_k+X_kA^{\mathrm{T}}+D+MX_kM^{\mathrm{T}}+(GX_kG^{\mathrm{T}})\circ(FX_kF^{\mathrm{T}})\|}{2\|A\|\|X_k\|+\|G\|^2\|F\|^2\|X_k\|^2+\|M\|^2\|X_k\|+\|D\|},$$

X_k represents the approximated solution when the algorithm stops at the tolerance $tol=10^{-12}$. The capital "IT" and "CPU" represent the number of iterations and the elapsed CPU time, respectively.

Example 1 Consider the QBEH (1) in [5] with

$$A=\begin{bmatrix}-2 & 1\\ 1 & -2\end{bmatrix},\quad G=I_2,\quad F=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix},$$
$$M=\begin{bmatrix}\sqrt{5/2} & 0\\ 0 & 0\end{bmatrix},\quad D=\begin{bmatrix}3 & -3\\ -3 & 3\end{bmatrix}.$$

This equation has a positive definite solution $\hat{X}=\mathrm{diag}(2,1)$. Starting from the zero matrix in our experiments, the sequences from the fixed-point iteration (4) and Newton's method (7) are both monotonically increasing and converge to the positive definite solution $\hat{X}$. The corresponding residual histories of the two methods are plotted in Fig. 1. One can see that both methods are efficient to calculate the solution to the QBEH. Specifically, the fixed-point method (4) requires 95 iterations (within 0.173 s) to attain the prescribed accuracy. Newton's method needs 5 iterations and the corresponding numbers of the fixed-point iteration for the ADI preconditioned subproblem are 43, 38, 27, 23, 23, respectively. The elapsed CPU time is about 0.188 s.

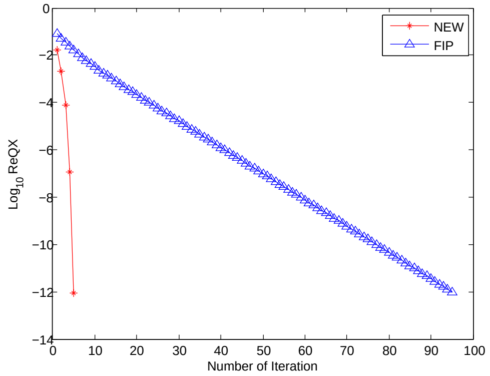

Fig. 1 Residual history in Example 1.

Example 2 This example is a proper modification of the transmission line circuit in [1, 3] as the original system is not stable. The coefficient matrices of the QBEH (1) are

$$A=\begin{bmatrix}A_{11} & A_{12}\\ A_{12}^{\mathrm{T}} & A_{22}\end{bmatrix},\quad D=\begin{bmatrix}D_{11} & D_{12}\\ D_{21} & D_{22}\end{bmatrix},$$
$$M=\begin{bmatrix}15.9107I_n & 0_{n\times n}\\ 0_{n\times n} & 0_{n\times n}\end{bmatrix},\quad G=I_N,\quad F=\begin{bmatrix}0_{n\times n} & 0_{n\times n}\\ -3I_n & -3I_n\end{bmatrix}$$

with $A_{11}=A_{22}=-18I_n$, $A_{12}=\mathrm{tridiag}(1,-3,1)$, $A_{12}(1,2)=A_{12}(2,1)=-1$; $D_{11}=D_{22}=0.0034I_n$, $D_{12}=D_{21}=-0.0137I_n$.

This equation was shown at least having a positive definite solution [5]. We took different $n's$ to test the fixed-point iteration (4) and Newton's iteration (7). Analogously to Example 1, sequences $\{X_k\}$ generated by both algorithms, starting from the initial zero matrix, were monotonically increasing and converge to the minimal positive solution. Table 1 records the "IT", "CPU" and "ReQX" of two algorithms at $n=20,40,60,80$.

It is seen from Table 1 that Newton's method always requires 10 iterations with each costing 8 ADI inner iterations to obtain the prescribed residual tolerance, beating the fixed-point method that needs 957, 995, 1001 and 1002 iterations at various n. Moreover, Newton's method spent less CPU time on attaining the terminated condition, indicating its superiority over the fixed-point iteration.

Example 3 This example is to show the effectiveness of the double Newton step when the QBEH is in the semi-stable case. The coefficient matrices of the QBEH (1) are

$$A=\begin{bmatrix}-2 & 1\\ 2 & -3\end{bmatrix},\quad D=5.543\begin{bmatrix}2.6141735 & -3\\ -3 & 3.6141735\end{bmatrix},$$
$$M=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix},\quad G=I_2,\quad F=0.5\times I_2.$$

This equation has a positive definite solution $\hat{X}=5.543\times I_2$. Moreover, $\mathscr{Q}'_{\hat{X}}$ has an eigenvalue of 4.8×10^{-5} and is almost semi-stable. Given the initial zero matrix, we ran the fixed-point iteration (4), Newton's method (7) and the double Newton's method (15) (DN-NEW) to compute the solution. In the DN-NEW, initial iterations in the double Newton's method are the same as Newton's iteration (7). But when the current iteration point is very close to the solution, i.e. the residual ReQX is less than 10^{-9},

Table 1 Numerical results for Example 2.

	Alg	IT	CPU	ReQX	Alg	IT	CPU	ReQX
$n = 20$	FIP	957	1.32	9.86×10^{-13}	NEW	10 (8)	0.703	9.84×10^{-13}
$n = 40$	FIP	995	4.55	9.97×10^{-13}	NEW	10 (8)	3.150	9.85×10^{-13}
$n = 60$	FIP	1001	8.48	9.94×10^{-13}	NEW	10 (8)	6.640	9.81×10^{-13}
$n = 80$	FIP	1002	13.47	9.94×10^{-13}	NEW	10 (8)	12.060	9.81×10^{-13}

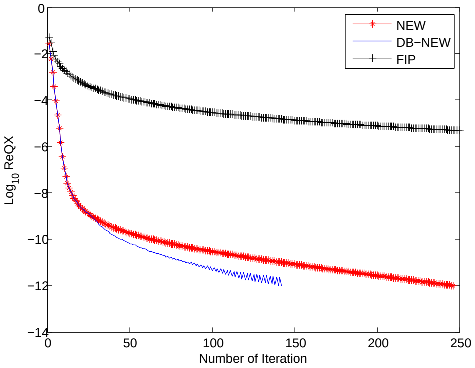

Fig. 2 Residual history for different methods in Example 3.

the double Newton step is employed. Fig. 2 records the residual history of different iterative methods. Note that the fixed-point method did not attain the prescribed residual level of 10^{-13} after 20 000 iterations (actually only arriving at 7.96×10^{-10}). So its residual history only for the first 250 iterations was plotted in contrast to Newton's method.

It is seen from Fig. 2 that Newton's method requires 246 iterations with each costing 11 ADI inner iterations to obtain the prescribed residual tolerance of 9.81×10^{-13}. The double Newton's method shares the same decreasing residual levels to Newton's method at about the former 30 steps, then drops lower at subsequent iterations. At last, it terminated after 142 iterations, arriving at the residual level of 9.66×10^{-13}. This shows that the double Newton steps (15) can effectively accelerate the iteration when the equation is in the semi-stable case.

Acknowledgements: The authors would like to thank the editor and two anonymous referees for their helpful comments to improve the original paper. This work was supported partly by NSF of China (11801163), NSF of Hunan Province (2020JJ4265, 2021JJ50032) and Key and General foundation of the Educational Department of Hunan Province (20A150, 20C0643).

REFERENCES

1. Benner P, Goyal P (2017) Balanced truncation model order reduction for quadratic-bilinear control systems. *arXiv:1705.00160v1*.
2. Antoulas AC (2005) *Approximation of Large-Scale Dynamical Systems*, SIAM, PA.
3. Gu C (2011) QLMOR: A projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems. *IEEE Trans Comput Aided Design Integr Circuits Sys* **30**, 1307–1320.
4. Schilders WHA, Van Der Vorst HA, Rommes J (2008) *Model Order Reduction: Theory, Research Aspects and Applications*, Springer-Verlag, Berlin.
5. Yu B, Dong N, Tang Q (2021) Existence of the solution to the quadratic bilinear equation arising from a class of quadratic dynamical systems. *arXiv:2107.03847*.
6. Wachspress EL (1995) *The ADI Model Problem*, Windsor, CA.
7. Wachspress EL (2013) Lyapunov and Sylevester matrix equations. In: *The ADI Model Problem*, Springer New York, pp 103–114.
8. Guo CH (2001) Iterative solution of a matrix Riccati equation arising in stochastic control. *Oper Theory Adv Appl* **130**, 209–221.
9. Guo CH (2001) Convergence rate of an iterative method for a nonlinear matrix equation. *SIAM J Matrix Anal Appl* **23**, 295–302.
10. Bhatia R (1997) *Matrix Analysis, Graduate Texts in Mathematics*, Springer, Berlin.
11. Schneider H (1968) Positive operators and an inertia theorem. *Numer Math* **7**, 11–17.
12. Damm T, Hinrichsen D (2001) Newton's method for a rational matrix equation occuring in stochastic control. *Linear Algebra Appl* **332–334**, 81–109.
13. Lancaster P, Rodman L (1995) *Algebraic Riccati Equations*, Clarendon Press, Oxford.
14. Ledermann W (1983) Issai Schur and his school in Berlin. *Bull London Math Soc* **15**, 97–106.
15. Horn RA, Mathias R (2001) An analog of the Cauchy-Schwarz inequality for Hadamard products and unitarily invariant norms. *SIAM J Matrix Anal Appl* **11**, 481–498.
16. Damm T (2008) Direct methods and ADI-preconditioned Krylov subspace methods for generalized Lyapunov equations. *Numer Linear Algebra Appl* **15**, 853–871.
17. Yu B, Li DH, Dong N (2013) Low memory and low complexity iterative schemes for a nonsymmetric algebraic Riccati equation arising from transport theory. *J Comput Appl Math* **250**, 175–189.