Proceedings of the Estonian Academy of Sciences,
2021, **70**, 3, 248–259
https://doi.org/10.3176/proc.2021.3.04
Available online at www.eap.ee/proceedings

COMPUTATIONAL
MATHEMATICS

A quadratic bilinear equation arising from the quadratic dynamical system

Bo Yu, Ning Dong* and Qiong Tang

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

Received 5 February 2021, accepted 9 June 2021, available online 4 August 2021

Abstract. A quadratic dynamical system with practical applications is taken into consideration. This system is transformed into a new bilinear system with Hadamard products by means of the implicit matrix structure. The corresponding quadratic bilinear equation is subsequently established via the Volterra series. Under proper conditions, the existence of the solution to the equation is proved by using a fixed-point iteration. Numerical experiments verify the proposed theory of the solution.

Key words: quadratic bilinear system, Kronecker product, Hadamard product, existence of the solution, fixed-point iteration.

1. INTRODUCTION

Consider a single-input and single-output quadratic dynamical system (QDS)

$$\begin{aligned} \dot{x}(t) &= Ax(t) + g(x(t), u(t)), \\ y(t) &= Cx(t), \end{aligned} \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state vector of time t, $u(t) \in \mathbb{R}$ denotes an input function, $g \in \mathbb{R}^n$ represents a quadratic function of $u(t)$ and $x(t)$, $y(t) \in \mathbb{R}$ is the output function, $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{1\times n}$ are the state and the output matrices, respectively. This system is one of the simplest nonlinear systems and is widely used in many applications [1,4,12,22,24]. Consider, for example, a transmission line circuit consisting of resistors, capacitors, and diodes with a constitutive nonlinear function $i_d(v) = e^{av} - 1, (a > 0)$ [4,12]. Assumed that, for simplicity, all resistors and capacitors have unit resistance and capacitance, then the input and output are the entering current source and the voltage at the first node, respectively. The corresponding differential system for this circuit at various nodes is

$$\begin{aligned} \dot{v}_1 &= -2v_1 + v_2 + 2 - e^{av_1} - e^{a(v_1 - v_2)} + u(t), \\ \dot{v}_i &= v_{i-1} - 2v_i + v_{i+1} + e^{a(v_{i-1} - v_i)} - e^{a(v_i - v_{i+1})}, \quad 2 \le i \le n-1, \\ \dot{v}_n &= v_{n-1} + v_n - 1 + e^{a(v_{n-1} - v_n)}. \end{aligned}$$

* Corresponding author, dongning_158@sina.com

To linearize the above nonlinear system, one can define variables $w_{i1} := e^{av_i}$ and $w_{i2} := e^{-av_i}$ to obtain a system of order at least $3n$. In contrast, another difference step might further reduce the order of the system. In fact, by setting $v_{i,i+1} = v_i - v_{i-1}$ as in [12], one has

$$\begin{aligned}
&\dot{v}_1 = -v_1 - v_{12} + 2 - e^{av_1} - e^{av_{12}} + u(t),\\
&\dot{v}_{12} = -v_1 - 2v_{12} + v_{23} + 2 - e^{av_1} - e^{av_{12}} + e^{av_{23}} + u(t),\\
&\dot{v}_{i,i+1} = v_{i-1,i} - 2v_{i,i+1} + v_{i+1,i+2} + e^{av_{i-1,i}} - 2e^{av_{i,i+1}} + e^{av_{i+1,i+2}}, \quad 2 \le i \le n-2,\\
&\dot{v}_{n-1,n} = v_{n-2,n-1} - 2v_{n-1,n} + 1 + e^{av_{n-2,n-1}} - 2e^{av_{n-1,n}}.
\end{aligned} \tag{2}$$

Let $w_1 = e^{av_1} - 1$ and $w_i = e^{av_{i-1,i}} - 1$ and differentiate both sides with respect to t. Then equations (2) can be further represented as

$$\begin{aligned}
&\dot{w}_1 = a(w_1+1)(-v_1 - v_{12} - w_1 - w_2 + u(t)),\\
&\dot{w}_2 = a(w_2+1)(-v_1 - 2v_{12} + v_{23} - w_1 - 2w_2 + w_3 + u(t)),\\
&\dot{w}_i = a(w_i+1)(v_{i-1,i} - 2v_{i,i+1} + v_{i+1,i+2} + w_{i-1} - 2y_i + y_{i+1}), \quad 2 \le i \le n-1,\\
&\dot{w}_n = a(w_n+1)(v_{n-2,n-1} - 2v_{n-1,n} + w_{n-1} - 2w_n).
\end{aligned} \tag{3}$$

Combining of (2) and (3) forms the quadratic bilinear system of order $N = 2n$ [4]

$$\begin{aligned}
&\dot{x}(t) = Ax(t) + H(x(t) \otimes x(t)) + Mx(t)u(t) + Bu(t),\\
&y(t) = Cx(t),
\end{aligned} \tag{4}$$

where the state vector is $x(t) = (\dot{v}_1, \dot{v}_{12} ..., \dot{v}_{n-1,n}, \dot{w}_1, ...\dot{w}_n)^\top \in \mathbb{R}^N$, the state matrix is

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \in \mathbb{R}^{N\times N}$$

with $A_i (i = 1,2,3,4)$ being the tri-diagonal matrix, $H \in \mathbb{R}^{N\times N^2}$ and $M \in \mathbb{R}^{N\times N}$ are sparse matrices associated with the quadratic functions $x(t) \otimes x(t)$ and $x(t)u(t)$, respectively, B is a vector of order N.

To efficiently control the quadratic system (4) when N is large, one has to search a low-dimensional (reduced-order) system to substitute for the original one, so that their systematic behaviours (for example, the stability and passivity) are sufficiently similar. Such a process is called the model order reduction (MOR) and has been well-established for linear systems in various areas [2,3]. One of the most popular MOR techniques is the balancing-type MOR, which has been successfully applied from the linear system to the nonlinear system [6,13]. This approach mainly relies on the controllability and the observability, or the Gramian matrix of the system, which is the solution to the corresponding algebraic matrix equation [4]:

$$AX + XA^\top + H(X \otimes X)H^\top + MXM^\top + D = 0 \tag{5}$$

with $D = BB^\top$. Obviously, solving the equation (5) involves a Kronecker product of the order N^2 and is normally expensive even if techniques of the truncation and compression [17] or the tensor matrization [19] are applied.

Noting the implicit structure in the original system, the system (4) can actually be transformed into another system to avoid the Kronecker product effectively. Indeed, let

$$F = \begin{bmatrix} 0_n & 0_n \\ A_3 & A_4 \end{bmatrix} \in \mathbb{R}^{N\times N}$$

and $G = I_N$. The quadratic item $H(x(t) \otimes x(t))$ in this example could be represented as $(Gx(t)) \circ (Fx(t))$, and thus the system (4) in [4] can be further rewritten as the quadratic bilinear system with the Hadamard product (QBSH)

$$\begin{aligned}
&\dot{x}(t) = Ax(t) + (Gx(t)) \circ (Fx(t)) + Mx(t)u(t) + Bu(t),\\
&y(t) = Cx(t).
\end{aligned} \tag{6}$$

The greatest advantage of the system (6) is that the nonlinear item depends merely on the Hadamard product, instead of the Kronecker product, between two vectors. Hence, the computational cost could be significantly reduced, especially for large N. If the afore-mentioned balancing-type MOR is used for the order reduction, two problems are still supposed to be addressed:

- What is the form of the algebraic equation corresponding to the QBSH (6)?
- Does the solution to the corresponding algebraic equation exist?

This paper will give positive answers to the above two questions. Specifically, we will make use of the Volterra series [23] to construct the corresponding quadratic bilinear equation of the QBSH (6) in Section 2. In Section 3, the existence of the solution to the equation will be demonstrated by a fixed-point iteration. Several numerical examples are listed in Section 4 to show the validity of the developed theory and Section 5 concludes the whole paper.

To proceed, the initial condition in the system (6) is assumed to be $x(0) = 0$. Throughout this paper, it is written $A \geq B$ ($A > B$) for symmetric matrices A and B if $A - B$ is a symmetric positive semidefinite (definite) matrix. $\sigma(A)$ and $\rho(A)$ denote here the spectrum and the spectral radius of the matrix A, respectively. The definition of the stability and several lemmas are also required in this research.

Definition 1 ([5]). *The matrix A is called stable (or semi-stable) if its spectrum lies in the left half of the complex plane (or the left half of the complex plane plus the imaginary axis), i.e. $\sigma(A) \in \mathbb{C}_{<}^{N\times N}$ (or $\sigma(A) \in \mathbb{C}_{\leq}^{N\times N}$).*

Lemma 1 ([7,18]). *Let the matrix $A \in \mathbb{R}^{N\times N}$ be stable in a linear system*

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t), \;\; x(0) = 0. \end{aligned}$$

The matrix $X = \int_0^\infty e^{At} BB^\top e^{A^\top t} dt$ is the solution to the Lyapunov equation

$$AX + XA^\top + D = 0$$

with $D = BB^\top$.

Lemma 2 ([15]). *Let the matrix $A \in \mathbb{R}^{N\times N}$ be stable and $B \in \mathbb{R}^{N\times N}$ be symmetric. Then the Lyapunov equation*

$$AX + XA^\top = B$$

has a unique symmetric solution X. Moreover, $X \geq 0$ if $B \leq 0$.

Lemma 3 ([16]). *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 $A \geq 0$ and $B \geq 0$, then $A \circ B \geq 0$. Moreover, $A \circ B > 0$ when A has no zero row.

2. THE ALGEBRAIC EQUATION CORRESPONDING TO QBSH

In this section, we concentrate on the reachability Gramian matrix of the QBSH (6) by using the Volterra series. It will show that the Gramian matrix is the solution to a quadratic bilinear equation with the Hadamard product (QBEH).

Only the continuous time-invariant QBSH (6) is considered and the discrete one can be derived analogously. It is known from [23,24] that the output of a nonlinear system in the Volterra series depends on the input of the system at all times and it could be expanded as

$$y(t) = h_0 + \sum_{n=1}^{N} \int_a^b \cdots \int_a^b h_n(t_1, \ldots, t_n) \prod_{j=1}^{n} x(t - t_j)\, dt_j.$$

The function $h_n(t_1, \ldots, t_n)$ is called the order-n Volterra kernel.

Proposition 1. *The state vector of the QBSH* (6) *can be formulated as*

$$\begin{aligned} x(t) &= \int_0^t e^{At_1}Bu_{t_1}(t)dt_1 + \int_0^t\int_0^{t-t_1} e^{At_1}Me^{At_2}Bu_{t_1t_2}(t)u_{t_1}(t)dt_1dt_2 \\ &\quad + \int_0^t\int_0^{t-t_1}\int_0^{t-t_1} e^{At_1}((Ge^{At_2}B)\circ(Fe^{At_3}B))u_{t_1t_2}(t)u_{t_1t_3}(t)dt_1dt_2dt_3 \\ &\quad + \int_0^t\int_0^{t-t_1}\int_0^{t-t_1-t_2} e^{At_1}Me^{At_2}Me^{At_3}Bu_{t_1t_2t_3}(t)u_{t_1t_2}(t)u_{t_1}(t)dt_1dt_2dt_3 + \ldots \end{aligned} \quad (7)$$

with $u_{t_1,\ldots t_k}(t) = u(t-t_1-\ldots-t_k)$ *and* $k \geq 1$.

Proof. As the first equation in (6) is a differential system, one can integrate from both sides with respect to t and get

$$x(t) = \int_0^t e^{At_1}Bu_{t_1}(t)dt_1 + \int_0^t e^{At_1}Mx_{t_1}(t)u_{t_1}(t)dt_1 + \int_0^t e^{At_1}((Gx_{t_1}(t))\circ(Fx_{t_1}(t)))dt_1 \quad (8)$$

with $x_{t_1}(t) = x(t-t_1)$. If the integrated upper bound is replaced by $t-t_1$, $x_{t_1}(t)$ can also be represented as

$$x_{t_1}(t) = \int_0^{t-t_1} e^{At_2}Bu_{t_1t_2}(t)dt_2 + \int_0^{t-t_1} e^{At_2}Mx_{t_1t_2}(t)u_{t_1t_2}(t)dt_1 + \int_0^{t-t_1} e^{At_2}((Gx_{t_1t_2}(t))\circ(Fx_{t_1t_2}(t)))dt_2 \quad (9)$$

with $x_{t_1t_2}(t) = x(t-t_1-t_2)$. By inserting (9) into (8), one has

$$\begin{aligned} x(t) &= \int_0^t e^{At_1}Bu_{t_1}(t)dt_1 + \int_0^t\int_0^{t-t_1} e^{At_1}Me^{At_2}Bu_{t_1t_2}(t)u_{t_1}(t)dt_1dt_2 \\ &\quad + \int_0^t\int_0^{t-t_1} e^{At_1}Me^{At_2}Mx_{t_1t_2}(t)u_{t_1t_2}(t)u_{t_1}(t)dt_1dt_2 \\ &\quad + \int_0^t\int_0^{t-t_1}\int_0^{t-t_1} e^{At_1}((Ge^{At_1}B)\circ(Fe^{At_1}B))u_{t_1t_2}(t)u_{t_1t_3}(t)dt_1dt_2dt_3 + O(\int\int\int\int). \end{aligned} \quad (10)$$

Again, noting

$$\begin{aligned} x_{t_1t_2}(t) &= \int_0^{t-t_1-t_2} e^{At_3}Bu_{t_1t_2t_3}(t)dt_3 + \int_0^{t-t_1-t_2} e^{At_3}Mx_{t_1t_2t_3}(t)u_{t_1t_2t_3}(t)dt_3 \\ &\quad + \int_0^{t-t_1-t_2} e^{At_3}((Gx_{t_1t_2t_3}(t))\circ(Fx_{t_1t_2t_3}(t)))dt_3 \end{aligned} \quad (11)$$

and inserting (11) into (10), the representation of $x(t)$ in (7) holds true after rearranging some items. □

The above proposition describes the Volterra expansion of the state vector $x(t)$, which is helpful for constructing the quadratic bilinear equation. To see this, let

$$\begin{aligned} L_1(t_1) &= e^{At_1}B, \\ L_2(t_1,t_2) &= e^{At_2}Me^{At_1}B \\ &:= e^{At_2}ML_1(t_1), \\ L_3(t_1,t_2,t_3) &= e^{At_3}[(GL_1(t_1))\circ(FL_1(t_2)),\ Me^{At_2}Me^{At_1}B] \\ &:= e^{At_3}[(GL_1(t_1))\circ(FL_1(t_2)),\ ML_2(t_1,t_2)], \\ &\ldots \\ L_k(t_1,\ldots,t_k) &:= e^{At_k}[(GL_1(t_1))\circ(FL_{k-2}(t_2,\ldots,t_{k-1})), \\ &\qquad (GL_2(t_1,t_2))\circ(FL_{k-3}(t_3,\ldots,t_{k-1})), \\ &\qquad \ldots, \\ &\qquad (GL_{k-2}(t_1,\ldots,t_{k-2}))\circ(FL_1(t_{k-1})),\ ML_{k-1}(t_1,\ldots,t_{k-1})] \end{aligned}$$

for $k > 3$. The following theorem reveals that the reachability Gramian matrix is the solution to a QBEH.

Theorem 1. *Let A be the stable matrix in the QBSH* (6)*. Define the reachability Gramian matrix*

$$X=\sum_{i=1}^{\infty}\Big(\int_0^{\infty}\dots\int_0^{\infty}L_i(t_1,\dots,t_i)L_i(t_1,\dots,t_i)^{\top}dt_1\dots dt_i\Big).$$

Then X satisfies the QBEH

$$\mathcal{Q}(X)=AX+XA^{\top}+D+MXM^{\top}+(GXG^{\top})\circ(FXF^{\top})=0. \tag{12}$$

Proof. Let

$$X_1=\int_0^{\infty}L_1(t_1)L_1(t_1)^{\top}dt_1:=\int_0^{\infty}e^{At_1}BB^{\top}e^{A^{\top}t_1}dt_1.$$

It follows from Lemma 1 that X_1 is the solution to the Lyapunov equation

$$AX_1+X_1A^{\top}+D=0 \tag{13}$$

with $D=BB^{\top}$. Next, consider the integration of order-2

$$\begin{aligned}X_2&=\int_0^{\infty}\int_0^{\infty}L_2(t_1,t_2)L_2(t_1,t_2)^{\top}dt_1dt_2\\&=\int_0^{\infty}\int_0^{\infty}e^{At_2}ML_1(t_1)L_1(t_1)^{\top}M^{\top}e^{A^{\top}t_1}dt_1dt_2\\&=\int_0^{\infty}e^{At_2}M\Big(\int_0^{\infty}L_1(t_1)L_1(t_1)^{\top}dt_1\Big)M^{\top}e^{A^{\top}t_1}dt_2\\&=\int_0^{\infty}e^{At_2}MX_1M^{\top}e^{A^{\top}t_1}dt_2.\end{aligned}$$

By using Lemma 1 again, X_2 is the solution to the following equation

$$AX_2+X_2A^{\top}+MX_1M^{\top}=0. \tag{14}$$

Proceeding with the integration for $i\geq 3$, one can get

$$\begin{aligned}X_i&=\int_0^{\infty}\dots\int_0^{\infty}L_i(t_1,\dots t_i)L_i(t_1,\dots,t_i)^{\top}dt_1\dots dt_i\\&=\int_0^{\infty}e^{At_i}\Big[\Big(\int_0^{\infty}GL_1L_1^{\top}G^{\top}dt_1\Big)\circ\Big(\int_0^{\infty}\dots\int_0^{\infty}FL_{i-2}L_{i-2}^{\top}F^{\top}dt_2\dots dt_{i-2}\Big)\\&\quad+\dots+\Big(\int_0^{\infty}\dots\int_0^{\infty}GL_{i-2}L_{i-2}^{\top}G^{\top}dt_1\dots dt_{i-2}\Big)\circ\Big(\int_0^{\infty}FL_1L_1^{\top}F^{\top}dt_{i-1}\Big)\\&\quad+M\Big(\int_0^{\infty}\dots\int_0^{\infty}L_{i-1}L_{i-1}^{\top}dt_1\dots dt_{i-1}\Big)M^{\top}\Big]e^{A^{\top}t_i}dt_i\\&=\int_0^{\infty}e^{At_i}\Big[(GX_1G^{\top})\circ(FX_{i-2}F^{\top})+\dots+(GX_{i-2}G^{\top})\circ(FX_1F^{\top})+MX_iM^{\top}\Big]e^{A^{\top}t_i}dt_i,\end{aligned}$$

in which we used the property $(v\circ u)(v\circ u)^{\top}=(vv^{\top})\circ(uu^{\top})$ with vectors u and v. By Lemma 1, X_i satisfies the equation

$$AX_i+X_iA^{\top}+(GX_1G^{\top})\circ(FX_{i-2}F^{\top})+\dots+(GX_{i-2}G^{\top})\circ(FX_1F^{\top})+MX_iM^{\top}=0. \tag{15}$$

Now, sum up the equations (13), (14) and (15) for $i\geq 3$. One has

$$A\Big(\sum_{i=1}^{\infty}X_i\Big)+\Big(\sum_{i=1}^{\infty}X_i\Big)A^{\top}+BB^{\top}+M\Big(\sum_{i=1}^{\infty}X_i\Big)M^{\top}+\Big(G\Big(\sum_{i=1}^{\infty}X_i\Big)G^{\top}\Big)\circ\Big(F\Big(\sum_{i=1}^{\infty}X_i\Big)F^{\top}\Big)=0,$$

which takes the form of the QBEH (12) by letting $X=\sum_{i=1}^{\infty}X_i$. $\square$

Remark. (1) As mentioned before, the computational complexity of the Hadamard product in the equation (12) is $O(N^2)$, compared with $O(N^4)$ of the Kronecker product in the equation (5). Even though the truncation and compression [17] or the tensor matrization technique [4,19] can reduce the complexity for large-scale sparse matrices in the case of the Kronecker product, the Hadamard product is still more effective in saving the flops counts, especially for dense and structured matrices (for example, the diagonal-plus-low-rank structure).

(2) As the Hadamard product can be represented as the sum of rank-one matrices (see Sec. 3.6 of [11]), the derived equation (12) can also be rewritten as a generalized stochastic or rational Riccati equation in [4,8,10,20]. Here we always use the Hadamard product for the convenience of describing the existence of the solution.

3. EXISTENCE OF THE SOLUTION TO QBEH

In this section, we will show the existence of the solution to the QBEH (12). Let $\mathscr{L}$ be a linear operator $\mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N}$ given by

$$\mathscr{L}(X) = AX + XA^\top.$$

Consider the iteration scheme

$$\mathscr{L}(X_{k+1}) = -(GX_kG^\top)\circ(FX_kF^\top) - MX_kM^\top - D \tag{16}$$

with an initial X_0. The following theorem shows the existence of the solution.

Theorem 2. *Let A be a stable matrix. Suppose that there is a positive (semi-)definite matrix Z to the inequality $\mathscr{Q}(Z) \geq 0$ and an initial matrix X_0 such that $X_0 \geq Z$ and $\mathscr{Q}(X_0) \leq 0$. Then the fixed-point iteration* (16) *produces a matrix sequence $\{X_k\}$ such that for $k \geq 0$*
(1) $X_k \geq X_{k+1}$, $X_k \geq Z$, $\mathscr{Q}(X_k) \leq 0$;
(2) $\lim_{k\to\infty} X_k = X^*$ *is a positive (semi-)definite solution to the QBEH* (12). *Especially, X^* is the maximal solution if X_0 is an upper bound for all solutions.*

Proof. The theorem is proved by induction applied to

$$X_i \geq X_{i+1}, \;\; X_i \geq Z, \;\; \mathscr{Q}(X_i) \leq 0, \;\; i \geq 0. \tag{17}$$

For $i = 0$, the assumption admits $X_0 \geq Z$ and $\mathscr{Q}(X_0) \leq 0$. It follows from (16) that

$$\begin{aligned} & A(X_1 - X_0) + (X_1 - X_0)A^\top \\ = \;& -(GX_0G^\top)\circ(FX_0F^\top) - MX_0M^\top - D - AX_0 - X_0A^\top \\ = \;& -\mathscr{Q}(X_0), \end{aligned}$$

implying $X_0 \geq X_1$ by the assumption and Lemma 2. Thus, (17) holds for $i = 0$.

Now, suppose that (17) is true for $i = k$. We next show that it is valid for $i = k+1$. In fact, it follows from the iteration (16) that

$$\begin{aligned} & A(X_{k+1} - Z) + (X_{k+1} - Z)A^\top \\ = \;& -(GX_kG^\top)\circ(FX_kF^\top) - MX_kM^\top - D - AZ - ZA^\top \\ = \;& -(G(X_k - Z)G^\top)\circ(FX_kF^\top) - (GX_kG^\top)\circ(F(X_k - Z)F^\top) - M(X_k - Z)M^\top - \mathscr{Q}(Z). \end{aligned}$$

As $\mathscr{Q}(Z) \geq 0$, $X_k - Z \geq 0$ and X_k is positive (semi-)definite from the induction assumption, it follows from Lemma 2 that the solution $X_{k+1} - Z$ of the above equation is unique and positive (semi-)definite, i.e. $X_{k+1} \geq$

Z. Moreover, the iteration (16) also indicates

$$
\begin{aligned}
& A(X_{k+1}-X_{k+2})+(X_{k+1}-X_{k+2})A^\top \\
= \; & -(GX_kG^\top)\circ(FX_kF^\top)-MX_kM^\top+(GX_{k+1}G^\top)\circ(FX_{k+1}F^\top)+MX_{k+1}M^\top \\
= \; & -(G(X_k-X_{k+1})G^\top)\circ(F(X_k-X_{k+1})F^\top) \\
& -(GX_kG^\top)\circ(FX_{k+1}F^\top)-(GX_{k+1}G^\top)\circ(FX_kF^\top)-M(X_k-X_{k+1})M^\top \\
\leq \; & -(G(X_k-X_{k+1})G^\top)\circ(F(X_k-X_{k+1})F^\top)-2(GZG^\top)\circ(FZF^\top)-M(X_k-Z)M^\top,
\end{aligned}
$$

where the inequality follows from the induction $X_k \geq Z$ and the proved fact $X_{k+1} \geq Z$. Consequently, the right hand side of the inequality is negative semi-definite and the inequality $X_{k+1} \geq X_{k+2}$ holds true by Lemma 3. Finally, the inequality

$$
\begin{aligned}
& \mathscr{Q}(X_{k+1}) \\
= \; & AX_{k+1}+X_{k+1}A^\top+(GX_{k+1}G^\top)\circ(FX_{k+1}F^\top)+MX_kM^\top+D \\
= \; & A(X_{k+1}-X_{k+2})+(X_{k+1}-X_{k+2})A^\top \\
\leq \; & 0
\end{aligned}
$$

shows that the induction assumption (17) holds for $i=k+1$. Then the sequence $\{X_k\}$ is well defined and has a limit $\lim_{k\to\infty} X_k = X^*$. Moreover, $X^* \geq Z$. Taking the limit from both sides of the iteration (16) indicates that X^* is the solution to the QBEH (12). Furthermore, X^* is the maximal solution when X_0 is the upper bound of all solutions. □

Remark. For the rational Riccati equations in [10,14,20], the stochastic term generally forms a positive operator, pushing against the stability. Then the condition of the stochastic stability is required to guarantee the existence of the solution. However, in the QBEH (12), the nonlinear item will form a negative operator when shifted to the right of the equation. Then Lemma 2 is applicable by the assumption on the stability of A. The following theorem further indicates the linear convergence of the sequence $\{X_k\}$ in the fixed-point iteration (16).

Theorem 3. *Let X^* be the solution to the QBEH and the sequence $\{X_k\}$ be produced by the iteration* (16)*. Let*

$$\mathscr{M}_{X^*}(\cdot)=M(\cdot)M^\top+(GX^*G^\top)\circ(F(\cdot)F^\top)+(G(\cdot)G^\top)\circ(FX^*F^\top)$$

be a linear operator at the solution X^*. If $\rho(\mathscr{L}^{-1}\mathscr{M}_{X^*})<1$, then

$$\limsup_{k\to\infty} \sqrt[k]{\|X_k-X^*\|} \leq \rho(\mathscr{L}^{-1}\mathscr{M}_{X^*})<1$$

with $\|\cdot\|$ any matrix norm.

Proof. Rewrite the iteration (16) as $X_{k+1}=\mathscr{F}(X_k)$ with the operator

$$\mathscr{F}(\cdot)=\mathscr{L}^{-1}(-M(\cdot)M^\top-(G(\cdot)G^\top)\circ(F(\cdot)F^\top)-D).$$

Then the Fréchet derivative of $\mathscr{F}$ at the solution X^* is

$$\mathscr{F}'_{X^*}(\Delta)=\mathscr{L}^{-1}(-M\Delta M^\top-(G\Delta G^\top)\circ(FX^*F^\top)-(GX^*G^\top)\circ(F\Delta F^\top)).$$

The conclusion is readily drawn from a classic theorem of fixed-point iteration such as in [21]. □

Remark. (1) The solver of the QBEH (12) determines the effectiveness of the balancing type MOR. Theorem 3 indicates that the convergence rate of the fixed-point iteration (16) is linear when $\rho(\mathscr{L}^{-1}\mathscr{M}_{X^*}) < 1$. If $\rho(\mathscr{L}^{-1}\mathscr{M}_{X^*}) = 1$, the convergence of the iteration (16) will degenerate to be sub-linear. In any case, acceleration of the iteration (16) should be further considered.

(2) The initial $X_0 \geq Z$ in Theorem 2 is similar to the one in [9]. Usually, it is not easy to validate the condition $\mathscr{Q}(X_0) \leq 0$. However, there is another easier way to select the initial matrix and this will be discussed in future work.

(3) The condition of the convergence in Theorem 3 is somewhat equivalent to the stochastic stability for stochastic rational Riccati equations. See [8,10,14,20] as well as references therein for more details.

4. NUMERICAL EXPERIMENTS

In this section, the existence of the solution to the QBEH (12) is validated by numerical experiments. The fixed-point iteration scheme (16) was coded by MATLAB 2014 and all examples were run on a laptop with Intel i3-3240 3.4GHz processor and 8GB RAM. The terminated condition for the fixed-point iteration was ReQX $< tol$ with

$$\text{ReQX} = \frac{\|AX + XA^\top + D + MXM^\top + (GXG^\top) \circ (FXF^\top)\|}{2\|A\|\|X_k\| + \|G\|^2\|F\|^2\|X_k\|^2 + \|M\|^2\|X_k\| + \|D\|}$$

and the tolerance $tol = 10^{-12}$. After termination, X_k was taken as the approximated solution to the QBEH (12).

Example 1. Consider the QBEH with

$$A = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}, \quad G = I_2, \quad F = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

$$M = \begin{pmatrix} \sqrt{5/2} & 0 \\ 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 3 & -3 \\ -3 & 3 \end{pmatrix}.$$

The initial iteration matrix is

$$X_0 = \begin{pmatrix} 2.75 & 0.40 \\ 0.40 & 2.00 \end{pmatrix}$$

and

$$\mathscr{Q}(X_0) = \begin{pmatrix} -0.3245 & 0.1500 \\ 0.1500 & -0.2000 \end{pmatrix}$$

is a negative definite matrix. Then the iteration sequence $\{X_k\}$, dictated by Theorem 2, is monotonically decreasing and converges to the positive definite solution

$$X^* = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

The history of the residual and the minimal eigenvalue of the difference $X_k - X_{k+1}$ (MESD) are plotted in Fig.1. It can be seen that the convergence rate is linear and $\{X_k\}$ is monotonically decreasing. In this example, X_{99} attains the prescribed residual level and can be regarded as an approximated solution to the QBEH (12).

Example 2. This example is a proper modification of the transmission line circuit in [4,12] as the original system is not stable. The coefficient matrices of the QBEH (12) are adapted to

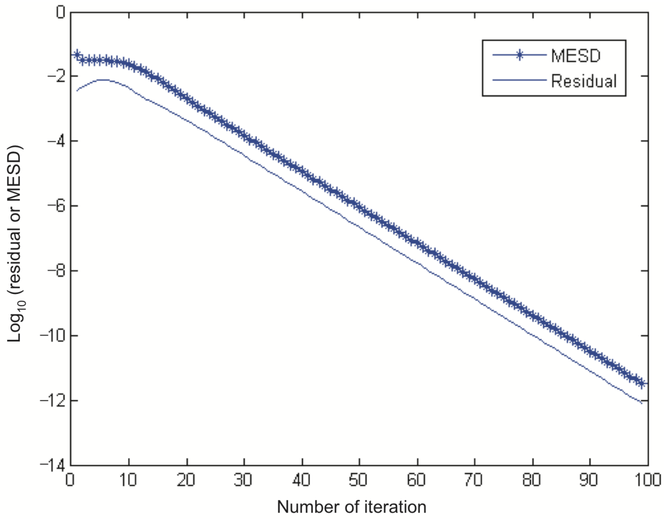

Fig. 1. Residual history for Example 1.

$$A=\begin{pmatrix} -18 & 0 & 0 & 0 & -3 & -1 & 0 & 0\\ 0 & -18 & 0 & 0 & -1 & -3 & 1 & 0\\ 0 & 0 & -18 & 0 & 0 & 1 & -3 & 1\\ 0 & 0 & 0 & -18 & 0 & 0 & 1 & -3\\ -3 & -1 & 0 & 0 & -18 & 0 & 0 & 0\\ -1 & -3 & 1 & 0 & 0 & -18 & 0 & 0\\ 0 & 1 & -3 & 1 & 0 & 0 & -18 & 0\\ 0 & 0 & 1 & -3 & 0 & 0 & 0 & -1 \end{pmatrix},$$

$$D=\begin{pmatrix} 0.0034 & 0 & 0 & 0 & -0.0137 & 0 & 0 & 0\\ 0 & 0.0034 & 0 & 0 & 0 & -0.0137 & 0 & 0\\ 0 & 0 & 0.0034 & 0 & 0 & 0 & -0.0137 & 0\\ 0 & 0 & 0 & 0.0034 & 0 & 0 & 0 & -0.0137\\ -0.0137 & 0 & 0 & 0 & 0.0549 & 0 & 0 & 0\\ 0 & -0.0137 & 0 & 0 & 0 & 0.0549 & 0 & 0\\ 0 & 0 & -0.0137 & 0 & 0 & 0 & 0.0549 & 0\\ 0 & 0 & 0 & -0.0137 & 0 & 0 & 0 & 0.0549 \end{pmatrix},$$

$$M=\begin{pmatrix} 15.9107I_4 & 0_{4\times4}\\ 0_{4\times4} & 0_{4\times4}\end{pmatrix},\ \ G=I_8,\ \ F=\begin{pmatrix} 0_{4\times4} & 0_{4\times4}\\ -3I_4 & -3I_4\end{pmatrix}.$$

Starting with

$$X_0=10^{-3}\begin{pmatrix} 8.5488 & 3.7695 & -1.3582 & 0.4268 & -1.3538 & -0.6580 & 0.2428 & -0.0778\\ 3.7695 & 9.9071 & -4.1963 & 1.3582 & -0.6580 & -1.5967 & 0.7358 & -0.2428\\ -1.3582 & -4.1963 & 9.9071 & -3.7695 & 0.2428 & 0.7358 & -1.5967 & 0.6580\\ 0.4268 & 1.3582 & -3.7695 & 8.5488 & -0.0778 & -0.2428 & 0.6580 & -1.3538\\ -1.3538 & -0.6580 & 0.2428 & -0.0778 & 1.7955 & 0.1985 & -0.0813 & 0.0264\\ -0.6580 & -1.5967 & 0.7358 & -0.2428 & 0.1985 & 1.8779 & -0.2246 & 0.0814\\ 0.2428 & 0.7358 & -1.5967 & 0.6580 & -0.0813 & -0.2246 & 1.8779 & -0.1982\\ -0.0778 & -0.2428 & 0.6580 & -1.3538 & 0.0264 & 0.0814 & -0.1982 & 1.7955 \end{pmatrix},$$

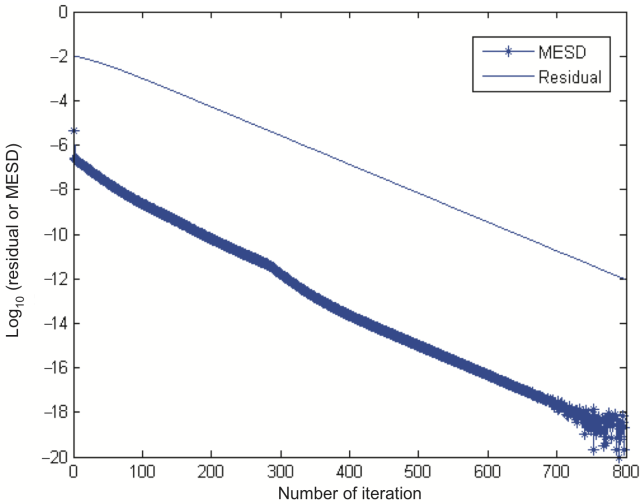

Fig. 2. Residual history for Example 2.

it is not difficult to see that $Q(X_0) \leq 0$. Similarly to the first example, the history of the residual and the minimal eigenvalue difference (MESD) are plotted in Fig. 2, indicating that the iteration sequence $\{X_k\}$ is monotonically decreasing and stops at $k = 798$ with the given tolerance $tol = 10^{-12}$. Accordingly,

$$X_{798} = 10^{-3} \begin{pmatrix} 4.3682 & 1.9044 & -0.6809 & 0.2135 & -0.9439 & -0.3574 & 0.1228 & -0.0390 \\ 1.9044 & 5.0492 & -2.1179 & 0.6809 & -0.3574 & -1.0667 & 0.3964 & -0.1228 \\ -0.6809 & -2.1179 & 5.0491 & -1.9044 & 0.1228 & 0.3964 & -1.0667 & 0.3574 \\ 0.2135 & 0.6809 & -1.9044 & 4.3681 & -0.0390 & -0.1228 & 0.3574 & -0.9439 \\ -0.9439 & -0.3574 & 0.1228 & -0.0390 & 1.7052 & 0.1189 & -0.0425 & 0.0133 \\ -0.3574 & -1.0667 & 0.3964 & -0.1228 & 0.1189 & 1.7479 & -0.1321 & 0.0425 \\ 0.1228 & 0.3964 & -1.0667 & 0.3574 & -0.0425 & -0.1321 & 1.7479 & -0.1188 \\ -0.0390 & -0.1228 & 0.3574 & -0.9439 & 0.0133 & 0.0425 & -0.1188 & 1.7052 \end{pmatrix}$$

can be taken as an approximation of X^*. It was also observed in our experiments that the residual level was dictated by Theorem 2 before the first 740 iterations, but fluctuated at subsequent iterations. Especially, the minimal eigenvalue of $X_k - X_{k+1}$ became negative at the steps of 745, 756, 765, 770, 774, 777, 778, 780, 782, 784, 785, 786, 790, 793 and 794. The reason might be that the current computational accuracy of the MESD possibly exceeds the machinery unit error $O(2^{-53})$. On the other hand, it also reflects that the fixed-point iteration (16) converges slowly at the neighbour of the true solution.

5. CONCLUSIONS

The quadratic bilinear system associated with the Kronecker product is rewritten as another system related to the Hadamard product according to the implicit matrix structure. The corresponding quadratic bilinear equation is subsequently obtained via the Volterra series and the existence of the solution is established by a fixed-point iteration. Several numerical experiments validate the proposed theoretical results. As the balancing type MOR method depends heavily on the solution to the QBEH (12), more efficient solvers might be developed in future research.

ACKNOWLEDGEMENTS

This work was supported partly by the NSF of China (11801163), NSF of Hunan Province (2020JJ4264, 2020JJ4265, 2021JJ50032), Key Foundation of the Education Department of Hunan Province (20A150) and General Foundation of the Education Department of Hunan Province (20C0643). The publication costs of this article were partially covered by the Estonian Academy of Sciences.

REFERENCES

1. Antoulas, A. C. *Approximation of Large-Scale Dynamical Systems*. SIAM Publications, Philadelphia, PA, 2005.
2. Astrid, P., Weiland, S., Willcox, K. and Backx, T. Missing point estimation in models described by proper orthogonal decomposition. *IEEE Trans. Automat. Control*, 2008, **53**(10), 2237–2251.
3. Barrault, M., Maday, Y., Nguyen, N. C. and Patera, A. T. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Math.*, 2004, **339**(9), 667–672.
4. Benner, P. and Goyal, P. Balanced truncation model order reduction for quadratic-bilinear control systems. 2017, *arXiv:1705.00160v1* [math.OC].
5. Bhatia, R. *Matrix Analysis, Graduate Texts in Mathematics*. Springer, Berlin, 1997.
6. Brennan, C., Condon, M. and Ivanov, R. Model order reduction of nonlinear dynamical systems. In *Progress in Industrial Mathematics at ECMI 2004* (Di Bucchianico, A., Mattheij, R. and Peletier, M., eds), vol. 8. Springer, Berlin, Heidelberg, 2006, 114–118.
7. Bubnicki, Z. *Modern Control Theory*. Springer, Berlin, Heidelberg, 2005.
8. Damm, T. *Rational Matrix Equations in Stochastic Control*. Springer, Berlin, Heidelberg, 2004.
9. Damm, T. and Hinrichsen, D. Newton's method for a rational matrix equation occurring in stochastic control. *Linear Algebra Appl.*, 2001, **332–334**(3), 81–109.
10. Fan, H.-Y., Weng P. C.-Y. and Chu, E. K.-W. Numerical solution to generalized Lyapunov/Stein and rational Riccati equations in stochastic control. *Numer. Algorithms*, 2016, **71**, 245–272.
11. Horn, R. A. and Johnson, C. R. The Hadamard product. In *Topics in Matrix Analysis*. Cambridge University Press, 1991, 298–381.
12. Gu, C. QLMOR: A projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.*, 2011, **30**(9), 1307–1320.
13. Gosea, V., Petreczky, M., Antoulas, A. C. and Fiter, C. Balanced truncation for linear switched systems. *Adv. Comput. Math.*, 2018, **44**(11), 1845–1886.
14. Guo, C.-H. Iterative solution of a matrix Riccati equation arising in stochastic control. *Oper. Theory Adv. Appl.*, 2001, **130**, 209–221.
15. Lancaster, P. and Rodman, L. *Algebraic Riccati Equations*. Clarendon Press, Oxford, 1995.
16. Ledermann, W. Issai Schur and his school in Berlin. *Bull. London Math. Soc.*, 1983, **15**(2), 97–106.
17. Li, T.-X., Weng, P. C.-Y., Chu, E. K.-W. and Lin, W.-W. Large-scale Stein and Lyapunov equations, Smith method, and applications. *Numer. Algorithms*, 2013, **63**(4), 727–752.
18. Ogata, K. *Modern Control Engineering.* Fifth edition. Pearson Education, Harlow, 2010.
19. Kolda, T. G. and Bader, B. W. Tensor decompositions and applications. *SIAM Rev.*, 2009, **51**(3), 455–500.
20. Ivanov, I. G. Iterations for solving a rational Riccati equation arising in stochastic control. *Comput. Math. Appl.*, 2007, **53**(6), 977–988.
21. Krasnoselskii, M. A., Vainikko, G. M., Zabreiko, P. P., Rutitskii, Ya. B. and Stetsenko, V. Ya. *Approximate Solution of Operator Equations*. Springer, Dordrecht, 1972.
22. Rewienski M. and White J. A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.*, 2003, **22**(2), 155–170.
23. Schetzen, M. *The Volterra and Wiener Theories of Nonlinear Systems*. Wiley-Interscience, New York, NY, 1980.
24. Schilders, W. H. A., van der Vorst, H. A. and Rommes, J. *Model Order Reduction: Theory, Research Aspects and Applications*. Springer, Berlin, Heidelberg, 2008.

Dünaamilisest ruutsüsteemist tulenev bilineaarne ruutvõrrand

Bo Yu, Ning Dong ja Qiong Tang

On vaadeldud praktilisi rakendusi omavat bilineaarsete ruutsüsteemide klassi. Esialgne süsteem on teisendatud Hadamardi korrutistega bilineaarseks süsteemiks ilmutamata maatriksstruktuuri abil. Vastav bilineaarne ruutvõrrand on arendatud ritta. Püsipunktiprintsiibi abil on tõestatud lahendi olemasolu ja esitatud numbriliste eksperimentide tulemusi.