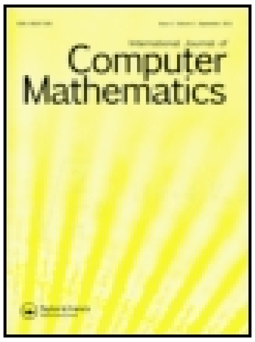

Taylor & Francis
Taylor & Francis Group

International Journal of Computer Mathematics

ISSN: 0020-7160 (Print) 1029-0265 (Online) Journal homepage: http://www.tandfonline.com/loi/gcom20

An iterative method for solving the continuous sylvester equation by emphasizing on the skew-hermitian parts of the coefficient matrices

Mohammad Khorsand Zak & Faezeh Toutounian

INTERNATIONAL JOURNAL OF COMPUTER MATHEMATICS, 2016
http://dx.doi.org/10.1080/00207160.2015.1120863

An iterative method for solving the continuous sylvester equation by emphasizing on the skew-hermitian parts of the coefficient matrices

Mohammad Khorsand Zak[a] and Faezeh Toutounian[b,c]

[a]Department of Applied Mathematics, Aligoudarz Branch, Islamic Azad University, Aligoudarz, Iran; [b]Department of Applied Mathematics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran; [c]The Center of Excellence on Modelling and Control Systems, Ferdowsi University of Mashhad, Mashhad, Iran

ABSTRACT

We present an iterative method based on the Hermitian and skew-Hermitian splitting (HSS) for solving the continuous Sylvester equation. By using the HSS of the coefficient matrices A and B, we establish a method which is practically inner/outer iterations, by employing a conjugate gradient on the normal equations (CGNR)-like method as inner iteration to approximate each outer iterate, while each outer iteration is induced by a convergent splitting of the coefficient matrices. Via this method, a Sylvester equation with coefficient matrices S_A and S_B (which are the skew-Hermitian part of A and B, respectively) is solved iteratively by a CGNR-like method. Convergence conditions of this method are studied and numerical examples show the efficiency of this method. In addition, we show that the quasi-Hermitian splitting can induce accurate, robust and effective preconditioned Krylov subspace methods.

ARTICLE HISTORY
Received 7 September 2015
Revised 29 October 2015
Accepted 6 November 2015

KEYWORDS
CGNR; Sylvester equation; Hermitian and skew-Hermitian splitting; preconditioning; nested iterations

2010 AMS SUBJECT CLASSIFICATIONS
15A24; 15A30; 15A69; 65F10; 65F30

1. Introduction

The Sylvester equation is ubiquitous in many areas of applied mathematics and plays vital roles in a number of applications such as control and system theory, model reduction and image processing, see [9–11] and references therein. The continuous Sylvester equation is possibly the most broadly employed linear matrix equation (see [2,8,9,11–13,18,20,22,23,26]), and is given as

$$AX + XB = C, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times m}$ and $C \in \mathbb{R}^{n\times m}$ are defined matrices and $X \in \mathbb{R}^{n\times m}$ is an unknown matrix. In general, the dimensions of A and B may be orders of magnitude different, and this fact is key in selecting the most appropriate numerical solution strategy. The continuous Sylvester equation (1) has a unique solution if and only if A and $-B$ have no common eigenvalues, which will be assumed throughout this paper.

Standard methods for numerical solution of the Sylvester equation (1) are the Bartels-Stewart [7] and the Hessenberg–Schur [15] methods, which consist of transforming coefficient matrices A and B into triangular or Hessenberg form by an orthogonal similarity transformation and then solving the resulting system directly by a back-substitution process. These methods are classified as direct

CONTACT Mohammad Khorsand Zak mo.khorsand@stu-mail.um.ac.ir

methods and are applicable and effective for general Sylvester equations of reasonably small size. When the coefficient matrices A and B are large and sparse, iterative methods such as the alternating direction implicit (ADI) method [8], the Krylov subspace based algorithms [13,18,19,26], the Hermitian and skew-Hermitian splitting (HSS) method [2,3], and the inexact variant of HSS (IHSS) iteration method [2,5] are often the methods of choice for efficiently and accurately solving the Sylvester equation (1). In [22,23], authors presented matrix variants of nested splitting conjugate gradient (NSCG) method which was first proposed in [1] for solving large sparse linear systems of equations. Recently, a preconditioned version of the NSCG method for generalized Sylvester equation is presented in [20].

In [23], the authors presented a class of NSCG method for the continuous Sylvester equation, in which both coefficient matrices are (non-Hermitian) positive semi-definite, and at least one of them is positive definite. This method is practically inner/outer iterations, which employs the Sylvester conjugate gradient method as inner iteration to approximate each outer iterate, while each outer iteration is induced by a convergent and Hermitian positive-definite splitting of the coefficient matrices. The NSCG method is suitable for problems with strong Hermitian part and it is not effective for problems with strong skew-Hermitian part [22,23].

In this paper, we present a new iterative method based on the HSS for solving the continuous Sylvester equation by emphasizing the role of skew-Hermitian part of the coefficient matrices. Similar to the NSCG method [23] by using the HSS of the coefficient matrices A and B, we establish a method which is inner/outer iterations, by employing a conjugate gradient on the normal equations (CGNR)-like method [32] as inner iteration to approximate each outer iterate, while each outer iteration is induced by a convergent splitting of the coefficient matrices. Via this method, which can abbreviate as NS-CGNR, a Sylvester equation with coefficient matrices S_A and S_B (which are the skew-Hermitian part of A and B, respectively) is solved iteratively by a CGNR-like method [32]. Convergence conditions of this method are studied and numerical experiments show the efficiency of this method. In addition, we show that the quasi-Hermitian splitting can induce accurate, robust and effective preconditioned Krylov subspace methods such as the BiCGSTAB method and the GMRES method.

The organization of this paper is as follows. Section 2 contains a brief preliminaries. Section 3 presents our own contribution, i.e., the NS-CGNR method for the continuous Sylvester equation and its convergence properties. Section 4 is devoted to numerical experiments. Finally, we present our conclusions in Section 5.

2. Preliminaries

In this section, we recall some necessary notations and useful results, which will be used in the following section. In this paper, we use $\lambda(M)$, $\|M\|_2$, $\|M\|_F$ and I_n to denote the eigenvalue, the spectral norm, the Frobenius norm of a matrix $M \in \mathbb{R}^{n\times n}$, and the identity matrix with dimension n, respectively. Note that $\|.\|_2$ is also used to represent the 2-norm of a vector. For nonsingular matrix $\mathcal{B}$, we denote by $\kappa(\mathcal{B}) = \|\mathcal{B}\|_2\|\mathcal{B}^{-1}\|_2$ its spectral condition number, and for a symmetric a positive-definite matrix $\mathcal{B}$, we define the $\|\cdot\|_{\mathcal{B}}$ norm of a vector $x \in \mathbb{R}^n$ as $\|x\|_{\mathcal{B}} = \sqrt{x^H\mathcal{B}x}$. Then the induced $\|\cdot\|_{\mathcal{B}}$ norm of a matrix $M \in \mathbb{R}^{n\times n}$ is define as $\|M\|_{\mathcal{B}} = \|\mathcal{B}^{1/2}M\mathcal{B}^{-1/2}\|_2$. In addition it holds that $\|Mx\|_{\mathcal{B}} \le \|M\|_{\mathcal{B}}\|x\|_{\mathcal{B}}$, $\|M\|_{\mathcal{B}} \le \sqrt{\kappa(\mathcal{B})}\|M\|_2$ and $\|I\|_{\mathcal{B}} = 1$, where I is the identity matrix. For any matrices $A = [a_{ij}]$ and $B = [b_{ij}]$, $A \otimes B$ denotes the Kronecker product defined as $A \otimes B = [a_{ij}B]$. For the matrix $X = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^{n\times m}$, $\mathrm{vec}(X)$ denotes the rvec operator defined as $\mathrm{vec}(X) = (x_1^{\mathrm{T}}, x_2^{\mathrm{T}}, \ldots, x_m^{\mathrm{T}})^{\mathrm{T}}$. Moreover, for a matrix $M \in \mathbb{R}^{n\times n}$ and the vector $\mathrm{vec}(M) \in \mathbb{R}^{nm}$, we have $\|M\|_F = \|\mathrm{vec}(M)\|_2$.

For matrix $\mathcal{A} \in \mathbb{R}^{n\times n}$, $\mathcal{A} = \mathcal{B} - \mathcal{C}$ is called a splitting of the matrix $\mathcal{A}$ if $\mathcal{B}$ is nonsingular. This splitting is a convergent splitting if $\rho(\mathcal{B}^{-1}\mathcal{C}) < 1$; and a contractive splitting if $\|\mathcal{B}^{-1}\mathcal{C}\| < 1$ for some matrix norm.

Now, consider the continuous Sylvester equation (1). It is mathematically equivalent to the linear system of equations

$$\mathcal{A}x = c, \tag{2}$$

where $x = \mathrm{vec}(X),\ c = \mathrm{vec}(C)$ and the matrix $\mathcal{A}$ is of dimension $nm \times nm$ and is given by

$$\mathcal{A} = I_m \otimes A + B^{\mathrm{T}} \otimes I_n. \tag{3}$$

Consider the HSS $\mathcal{A} = \mathcal{H} + \mathcal{S}$, where

$$\mathcal{H} = \frac{\mathcal{A} + \mathcal{A}^H}{2}, \quad \mathcal{S} = \frac{\mathcal{A} - \mathcal{A}^H}{2}, \tag{4}$$

are the Hermitian and skew-Hermitian parts of matrix $\mathcal{A}$, respectively, see [3,4]. Since the matrix $\mathcal{S}$ may be singular, we introduce a shift ($\alpha > 0$) and define quasi-Hermitian splitting

$$\mathcal{A} = (\mathcal{H} - \alpha I) + (\mathcal{S} + \alpha I) = \mathcal{H}_\alpha + \mathcal{S}_\alpha. \tag{5}$$

Then the system of linear equations (2) is equivalent to the fixed-point equation

$$\mathcal{S}_\alpha x = c - \mathcal{H}_\alpha x.$$

Now, with given an initial guess $x^{(0)} \in \mathbb{R}^n$, assume that we have computed approximations $x^{(1)}, x^{(2)}, \ldots, x^{(l)}$ to the solution $x^* \in \mathbb{R}^n$ of the system (2). Therefore, the next approximation $x^{(l+1)}$ may be defined as either an exact or an inexact solution of the system of linear equations

$$\mathcal{S}_\alpha x = c - \mathcal{H}_\alpha x^{(l)}. \tag{6}$$

Now, we can solve the linear system of equations (6) by the CGNR method [27]. Similar to [1], we can establish and prove the following theorem about the convergence properties of this method which can named as NS-CGNR method.

Theorem 2.1: *Let $\mathcal{A} \in \mathbb{R}^{nm\times nm}$ be a nonsingular and non-symmetric matrix, and $\mathcal{A} = \mathcal{H}_\alpha + \mathcal{S}_\alpha$ a contractive (with respect to the $\|\cdot\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}$-norm). Suppose that the NS-CGNR method is started from an initial guess $x^{(0)} \in \mathbb{R}^{nm}$, and produces an iterative sequence $\{x^{(l)}\}_{l=0}^{\infty}$, where $x^{(l)} \in \mathbb{R}^{nm}$ is the* l*th approximation to the solution $x^* \in \mathbb{R}^{nm}$ of the system of linear equations* (2), *obtained by solving the linear system* (6) *with k_l steps of CGNR iterations. Then*

(a) $\|x^{(l)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \le \gamma^{(l)}\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha},\ l = 1, 2, 3, \ldots,$
(b) $\|c - \mathcal{A}x^{(l)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \le \tilde{\gamma}^{(l)}\|c - \mathcal{A}x^{(l-1)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha},\ l = 1, 2, 3, \ldots,$

where

$$\gamma^{(l)} = 2\left(\frac{\kappa(\mathcal{S}_\alpha) - 1}{\kappa(\mathcal{S}_\alpha) + 1}\right)^{k_l}(1 + \varrho) + \varrho, \quad \tilde{\gamma}^{(l)} = \gamma^{(l)}\kappa(\mathcal{S}_\alpha)\frac{1 + \varrho}{1 - \varrho}, \quad l = 1, 2, 3, \ldots$$

and $\varrho = \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} = \|\mathcal{H}_\alpha\mathcal{S}_\alpha^{-1}\|_2$.

Moreover, for some $\gamma \in (\varrho, \varrho_1)$ with $\varrho_1 = \min\{1, 2 + 3\varrho\}$, and

$$k_l \ge \frac{\ln((\gamma - \varrho)/(2(1 + \varrho)))}{\ln((\kappa(\mathcal{S}_\alpha) - 1)/(\kappa(\mathcal{S}_\alpha) + 1))}, \quad l = 1, 2, 3, \ldots,$$

we have $\gamma^{(l)} \le \gamma (l = 1, 2, 3, \ldots)$, and the sequence $\{x^{(l)}\}_{l=0}^{\infty}$ converges to the solution x^* of the system of linear equations (2). For $\varrho \in (0, r)$, in which r *is the positive root of quadratic equation $\kappa(\mathcal{S}_\alpha)\varrho^2 +$*

$(\kappa(\mathcal{S}_\alpha)+1)\varrho-1=0$, *and some* $\tilde{\gamma}\in((1+\varrho)\varrho\kappa(\mathcal{S}_\alpha)/(1-\varrho),1)$, *and*

$$k_l \geq \frac{\ln(((1-\varrho)\tilde{\gamma}-\varrho(1+\varrho)\kappa(\mathcal{S}_\alpha))/(2(1+\varrho)^2\kappa(\mathcal{S}_\alpha)))}{\ln((\kappa(\mathcal{S}_\alpha)-1)/(\kappa(\mathcal{S}_\alpha)+1))}, \quad l=1,2,3,\ldots,$$

we have $\tilde{\gamma}^{(l)}\leq\tilde{\gamma}(l=1,2,3,\ldots)$, *and the residual sequence* $\{c-\mathcal{A}x^{(l)}\}_{l=0}^{\infty}$ *converges to zero.*

Proof: See Appendix. ■

From the work of Golub and Vandrestraeten [16], If

$$\lambda_{\min}(\mathcal{H})\lambda_{\max}(\mathcal{H}) > \min_{\lambda\in\Lambda(\mathcal{S})}|\lambda(\mathcal{S})|, \tag{7}$$

then there exists an α for which $\rho(\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha)<1$. Moreover, using

$$\alpha=\frac{\lambda_{\min}(\mathcal{H})+\lambda_{\max}(\mathcal{H})}{2}, \tag{8}$$

can cause to decrease the upper bound of $\rho(\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha)$, see [16].

3. The NS-CGNR method for the Sylvester equation

In this section, we establish the NS-CGNR method for solving the Sylvester equation (1). We suppose that both coefficient matrices in Equation (1) are (non-Hermitian) positive semi-definite, and at least one of them is positive definite. Consider the quasi-HSS $\mathcal{A}=\mathcal{H}_\alpha+\mathcal{S}_\alpha$. From Equations (3) and (4), by using the Kronecker product's properties [17,25], we have

$$\mathcal{H}_\alpha = I_m\otimes H_A(\alpha)+H_{B^{\mathrm{T}}}(\alpha)\otimes I_n, \tag{9}$$

$$\mathcal{S}_\alpha = I_m\otimes S_A(\alpha)+S_{B^{\mathrm{T}}}(\alpha)\otimes I_n, \tag{10}$$

where $H_A(\alpha)=H_A-\alpha I_n$, $S_A(\alpha)=S_A+\alpha I_n$, $H_{B^{\mathrm{T}}}(\alpha)=H_{B^{\mathrm{T}}}-\alpha I_m$, $S_{B^{\mathrm{T}}}(\alpha)=S_{B^{\mathrm{T}}}+\alpha I_m$ and H_A, S_A, $H_{B^{\mathrm{T}}}$, $S_{B^{\mathrm{T}}}$ are the Hermitian and skew-Hermitian parts of A and B^{T}, respectively. By using the relations (9) and (10), we obtain the following linear system of equations:

$$(I_m\otimes S_A(\alpha)+S_{B^{\mathrm{T}}}(\alpha)\otimes I_n)x = c-(I_m\otimes H_A(\alpha)+H_{B^{\mathrm{T}}}(\alpha)\otimes I_n)x^{(l)}, \tag{11}$$

which can be arranged equivalently as

$$S_A(\alpha)X+XS_B(\alpha)=C-H_A(\alpha)X^{(l)}-X^{(l)}H_B(\alpha). \tag{12}$$

The eigenvalues of a skew-Hermitian matrix are pure imaginary. So, we have $\mathrm{Re}(\lambda(S_A(\alpha)))=\alpha$ and $\mathrm{Re}(\lambda(-S_B(\alpha)))=-\alpha$. Due to this, we can easily see that there is no common eigenvalue between the matrices $S_A(\alpha)$ and $-S_B(\alpha)$, so the Sylvester equation (12) has unique solution for all given right-hand side matrices. For obtaining $X^{(l+1)}$, we can solve the matrix equation (12) iteratively by the Sylvester version of the CGNR method. Now, based on the above observations, we can establish the following algorithm for the NS-CGNR method for solving the continuous Sylvester equation (1).

3.1. Implementation of the NS-CGNR method

An implementation of the NS-CGNR method is given by the following algorithm. In the following algorithm, $k_{\max}$ and $j_{\max}$ are the largest admissible number of the outer and the inner iteration steps, respectively. $X^{(0)}$ is an initial guess for the solution, and the outer and the inner stopping tolerances are denoted by ϵ and η, respectively.

Algorithm 3.1 The NS-CGNR algorithm for the continuous Sylvester equation

1. $X^{(0,0)} = X^{(0)}$
2. $R^{(0)} = C - AX^{(0)} - X^{(0)}B$
3. For $k = 0, 1, 2, \ldots, k_{\max}$ Do:
4. $\quad \hat{C} = C - H_A(\alpha)X^{(k,0)} - X^{(k,0)}H_B(\alpha)$
5. $\quad \hat{R}^{(0)} = \hat{C} - S_A(\alpha)X^{(k,0)} - X^{(k,0)}S_B(\alpha)$
6. $\quad Z^{(0)} = S_A^{\mathrm{T}}(\alpha)\hat{R}^{(0)} + \hat{R}^{(0)}S_B^{\mathrm{T}}(\alpha)$, and $P^{(0)} = Z^{(0)}$
7. $\quad$ For $j = 0, 1, 2, \ldots, j_{\max}$ Do:
8. $\qquad W^{(j)} = S_A(\alpha)P^{(j)} + P^{(j)}S_B(\alpha)$
9. $\qquad \alpha_j = \frac{\|Z^{(j)}\|_F^2}{\|W^{(j)}\|_F^2}$
10. $\qquad X^{(k,j+1)} = X^{(k,j)} + \alpha_j P^{(j)}$
11. $\qquad \hat{R}^{(j+1)} = \hat{R}^{(j)} - \alpha_j W^{(j)}$
12. $\qquad$ If $\|\hat{R}^{(j+1)}\|_F \leq \eta\|\hat{R}^{(0)}\|_F$ GoTo 17
13. $\qquad Z^{(j+1)} = S_A^{\mathrm{T}}(\alpha)\hat{R}^{(j+1)} + \hat{R}^{(j+1)}S_B^{\mathrm{T}}(\alpha)$
14. $\qquad \beta_j = \frac{\|Z^{(j+1)}\|_F^2}{\|Z^{(j)}\|_F^2}$
15. $\qquad P^{(j+1)} = Z^{(j+1)} + \beta_j P^{(j)}$
16. $\quad$ End Do
17. $\quad X^{(k+1)} = X^{(k,j)}$
18. $\quad R^{(k+1)} = C - AX^{(k+1)} - X^{(k+1)}B$
19. $\quad$ If $\|R^{(k+1)}\|_F \leq \epsilon\|R^{(0)}\|_F$ Stop
20. $\quad X^{(k+1,0)} = X^{(k+1)}$
21. End Do

3.2. Convergence analysis

In the sequel, we need the following lemmas.

Lemma 3.2 ([21]): *Let $\mathcal{A} \in \mathbb{R}^{n\times n}$ be a symmetric positive-definite matrix. Then for all $x \in \mathbb{R}^n$, we have $\|\mathcal{A}^{1/2}x\|_2 = \|x\|_{\mathcal{A}}$ and*

$$\sqrt{\lambda_{\min}(\mathcal{A})}\|x\|_{\mathcal{A}} \leq \|\mathcal{A}x\|_2 \leq \sqrt{\lambda_{\max}(\mathcal{A})}\|x\|_{\mathcal{A}}.$$

Lemma 3.3 ([25]): *Suppose that $A, B \in \mathbb{R}^{n\times n}$ be two Hermitian matrices, and denote the minimum and the maximum eigenvalues of a matrix M with $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$, respectively. Then*

$$\lambda_{\max}(A + B) \leq \lambda_{\max}(A) + \lambda_{\max}(B),$$
$$\lambda_{\min}(A + B) \geq \lambda_{\min}(A) + \lambda_{\min}(B).$$

Lemma 3.4: *For matrix $\mathcal{S}_\alpha = I_m \otimes S_A(\alpha) + S_{B^{\mathrm{T}}}(\alpha) \otimes I_n$, we have*

$$\|\mathcal{S}_\alpha\|_2 \leq \max|\lambda(S_A(\alpha))| + \max|\lambda(S_B(\alpha))|.$$

Proof: For skew-Hermitian matrix S_A, we have

$$\begin{aligned}\|S_A(\alpha)\|_2^2 &= \lambda_{\max}(S_A(\alpha)^H S_A(\alpha))\\ &= \lambda_{\max}((S_A + \alpha I)^H(S_A + \alpha I))\\ &= \lambda_{\max}(S_A^H S_A + \alpha^2 I)\\ &= |\lambda_{\max}(S_A + \alpha I)|^2.\end{aligned}$$

Therefore, $\|S_A(\alpha)\|_2 = |\lambda_{\max}(S_A + \alpha I)|$ and $\|S_B(\alpha)\|_2 = |\lambda_{\max}(S_B + \alpha I)|$. Moreover, we have

$$\begin{aligned}\|\mathcal{S}(\alpha)\|_2 &= \|I_m \otimes S_A(\alpha) + S_B^{\mathrm{T}}(\alpha) \otimes I_n\|_2 \\ &\leq \|I_m \otimes S_A(\alpha)\|_2 + \|S_B^{\mathrm{T}}(\alpha) \otimes I_n\|_2 \\ &= \|S_A(\alpha)\|_2 + \|S_B^{\mathrm{T}}(\alpha)\|_2.\end{aligned}$$

Therefore, $\|\mathcal{S}_\alpha\|_2 \leq \max|\lambda(S_A(\alpha))| + \max|\lambda(S_B(\alpha))|$. ■

Lemma 3.5 ([25]): *Let $A, B \in \mathbb{R}^{n\times n}$, and λ and μ be the eigenvalues of* A *and* B, *and* x *and* y *be the corresponding eigenvectors, respectively. Then $\lambda\mu$ is an eigenvalue of $A \otimes B$ corresponding to the eigenvector $x \otimes y$.*

For the linear system of Equations (2) when the coefficient matrix is non-Hermitian positive definite, Wang and Bai [29] presented sufficient conditions for the convergent splittings. Also, the convergence conditions for splitting iteration methods for non-Hermitian coefficient matrix are studied in [30]. Some similar results were presented in [22,23] in the case of matrix equations. In the following lemma, we state and prove a result for splitting (5) in the case of matrix equations for the Sylvester equation (1).

Lemma 3.6: *Suppose that $\mathcal{H}_\alpha$ and $\mathcal{S}_\alpha$ are as in Equations* (9) *and* (10), *respectively. If*

$$\frac{\lambda_{\max}(H_A(\alpha)) + \lambda_{\max}(H_B(\alpha))}{\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|} < \frac{1}{\theta^4},$$

where $\theta = (\max|\lambda(S_A(\alpha))| + \max|\lambda(S_B(\alpha))|)/(\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|)$, then $\mathcal{A} = \mathcal{H}_\alpha + \mathcal{S}_\alpha$ is a contractive splitting (with respect to the $\|\cdot\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}$-norm), i.e., $\|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} < 1$.

Proof: From Equations (9) and (10), by Lemmas 3.2–3.4, we have

$$\|\mathcal{H}_\alpha\|_2 = \lambda_{\max}(\mathcal{H}_\alpha) \leq \lambda_{\max}(H_A(\alpha)) + \lambda_{\max}(H_B(\alpha)),$$

and

$$\|\mathcal{S}_\alpha\|_2 \leq \max|\lambda(S_A(\alpha))| + \max|\lambda(S_B(\alpha))|,$$

where $\Lambda(\mathcal{S}_\alpha)$ is the set of all eigenvalues of $\mathcal{S}_\alpha$. Therefore, it follows that

$$\begin{aligned}\|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} &\leq \sqrt{\kappa(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)}\|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_2 \\ &= \sqrt{\kappa(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)}\|(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)^{-1}\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{H}_\alpha\|_2 \\ &\leq \sqrt{\kappa(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)}\|(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)^{-1}\|_2\|\mathcal{S}_\alpha^{\mathrm{T}}\|_2\|\mathcal{H}_\alpha\|_2.\end{aligned}$$

Therefore, we have

$$\|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \leq (\kappa(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha))^{3/2}\frac{\|\mathcal{S}_\alpha^{\mathrm{T}}\|_2\|\mathcal{H}_\alpha\|_2}{\|\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha\|_2}. \tag{13}$$

Again, the use of Lemmas 3.2 and 3.4 implies that

$$\begin{aligned}\kappa(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha) &= \frac{\lambda_{\max}(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)}{\lambda_{\min}(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)} \\ &= \frac{\max_{\lambda\in\Lambda(\mathcal{S}_\alpha)}|\lambda(\mathcal{S}_\alpha)|^2}{\min_{\lambda\in\Lambda(\mathcal{S}_\alpha)}|\lambda(\mathcal{S}_\alpha)|^2} \\ &\leq \left(\frac{\max|\lambda(S_A(\alpha))| + \max|\lambda(S_B(\alpha))|}{\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|}\right)^2.\end{aligned}$$

So, we can write

$$\sqrt{\kappa(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)} \leq \frac{\max|\lambda(S_A(\alpha))| + \max|\lambda(S_B(\alpha))|}{\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|} = \theta. \tag{14}$$

Moreover, we have

$$\begin{aligned} \|\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha\|_2 &= \max_{\lambda\in\Lambda(\mathcal{S}_\alpha)} |\lambda(\mathcal{S}_\alpha)|^2 \\ &\geq \min_{\lambda\in\Lambda(\mathcal{S}_\alpha)} |\lambda(\mathcal{S}_\alpha)|^2 \\ &\geq (\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|)^2. \end{aligned} \tag{15}$$

Therefore,

$$\begin{aligned} \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} &\leq \left(\frac{\max|\lambda(S_A(\alpha))| + \max|\lambda(S_B(\alpha))|}{\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|}\right)^4 \frac{\lambda_{\max}(H_A(\alpha)) + \lambda_{\max}(H_B(\alpha))}{\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|} \\ &= \theta^4 \frac{\lambda_{\max}(H_A(\alpha)) + \lambda_{\max}(H_B(\alpha))}{\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|}. \end{aligned} \tag{16}$$

This clearly proves the lemma. ■

Let

$$\eta = \theta^4 \frac{\lambda_{\max}(H_A(\alpha)) + \lambda_{\max}(H_B(\alpha))}{\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|},$$

and $\varrho = \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}$. When $\eta < 1$, from the proof of Lemma 3.5 (relation (16)), we have $\varrho \leq \eta < 1$. Therefore, $\mathcal{A} = \mathcal{H}_\alpha + \mathcal{S}_\alpha$ is a contractive (in the $\|\cdot\|_{\mathcal{S}^{\mathrm{T}}\mathcal{S}}$ norm). In this case, Lemma 3.1 and part (a) of Theorem 2.1 imply that

$$\begin{aligned} \frac{1}{\sqrt{\lambda_{\max}(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)}}\|\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha(x^{(l)} - x^*)\|_2 &\leq \|x^{(l)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\ &\leq \gamma^{(l)}\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\ &\leq \frac{\gamma^{(l)}}{\sqrt{\lambda_{\min}(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)}}\|\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha(x^{(l-1)} - x^*)\|_2, \end{aligned}$$

where

$$\gamma^{(l)} = 2\left(\frac{\kappa(\mathcal{S}_\alpha) - 1}{\kappa(\mathcal{S}_\alpha) + 1}\right)^{k_l}(1 + \varrho) + \varrho.$$

So, part (a) in Theorem 2.1 can be written as

$$\|\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha(x^{(l)} - x^*)\|_2 \leq \gamma^{(l)}\frac{\sqrt{\lambda_{\max}(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)}}{\sqrt{\lambda_{\min}(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)}}\|\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha(x^{(l-1)} - x^*)\|_2. \tag{17}$$

Furthermore, from Equation (14)

$$\gamma^{(l)} \leq 2\left(\frac{\theta - 1}{\theta + 1}\right)^{k_l}(1 + \eta) + \eta.$$

Now, by using

$$\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha = I_m \otimes S_A^{\mathrm{T}}(\alpha)S_A(\alpha) + S_{B^{\mathrm{T}}}(\alpha) \otimes S_A^{\mathrm{T}}(\alpha) + S_{B^{\mathrm{T}}}^{\mathrm{T}}(\alpha) \otimes S_A(\alpha) + S_{B^{\mathrm{T}}}^{\mathrm{T}}(\alpha)S_{B^{\mathrm{T}}}(\alpha) \otimes I_n,$$

and notation

$$\hat{E}^{(l)} = S_A^{\mathrm{T}}(\alpha)S_A(\alpha)E^{(l)} + S_A^{\mathrm{T}}(\alpha)E^{(l)}S_{B^{\mathrm{T}}}^{\mathrm{T}}(\alpha) + S_A(\alpha)E^{(l)}S_{B^{\mathrm{T}}}(\alpha) + E^{(l)}S_{B^{\mathrm{T}}}^{\mathrm{T}}(\alpha)S_{B^{\mathrm{T}}}(\alpha), \tag{18}$$

where $E^{(l)} = X^{(l)} - X^*$, the relation (17) can be arranged equivalently as

$$\|\hat{E}^{(l)}\|_F \le \omega^{(l)}\|\hat{E}^{(l-1)}\|_F, \tag{19}$$

where

$$\omega^{(l)} = \left(2\left(\frac{\theta-1}{\theta+1}\right)^{k_l}(1+\eta)+\eta\right)\theta.$$

It is obvious that, for $\eta \in (0, 1/\theta)$ and $\omega \in (\eta\theta, 1)$, we will have $\omega^{(l)} \le \omega$ if

$$k_l \ge \frac{\ln((\omega-\eta\theta)/2\theta(1+\eta))}{\ln((\theta-1)/(\theta+1))}, \quad l = 1, 2, 3, \ldots,$$

Under this restriction, from Equation (19), we have

$$\begin{aligned}\|\hat{E}^{(l)}\|_F &\le \omega^{(l)}\|\hat{E}^{(l-1)}\|_F \\ &\le \Pi_{k=0}^{l}\omega^{(k)}\|\hat{E}^{(0)}\|_F \\ &\le \omega^{l+1}\|\hat{E}^{(0)}\|_F.\end{aligned}$$

Therefore, the sequence $\{X^{(l)}\}_{l=0}^{\infty}$ converges to the solution X^* of the system of linear equations (1).

Similarly, by using part (b) of Theorem 2.1, for residual $r^{(l)} = c - \mathcal{A}x^{(l)}$, we obtain

$$\|\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha r^{(l)}\|_2 \le \tilde{\gamma}^{(l)}\frac{\sqrt{\lambda_{\max}(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)}}{\sqrt{\lambda_{\min}(\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha)}}\|\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha r^{(l-1)}\|_2, \tag{20}$$

where $\tilde{\gamma}^{(l)} = \gamma^{(l)}\kappa(\mathcal{S}_\alpha)(1+\varrho)/(1-\varrho)$, and

$$\tilde{\gamma}^{(l)} \le \theta\left(2\left(\frac{\theta-1}{\theta+1}\right)^{k_l}(1+\eta)+\eta\right)\frac{1+\eta}{1-\eta}.$$

Now, by using notation

$$\hat{R}^{(l)} = S_A^{\mathrm{T}}(\alpha)S_A(\alpha)R^{(l)} + S_A^{\mathrm{T}}(\alpha)R^{(l)}S_{B^{\mathrm{T}}}^{\mathrm{T}}(\alpha) + S_A(\alpha)R^{(l)}S_{B^{\mathrm{T}}}(\alpha) + R^{(l)}S_{B^{\mathrm{T}}}^{\mathrm{T}}(\alpha)S_{B^{\mathrm{T}}}(\alpha), \tag{21}$$

where $R^{(l)} = C - AX^{(l)} - X^{(l)}B$, the relation (20) can also be arranged equivalently as

$$\|\hat{R}^{(l)}\|_F \le \tilde{\omega}^{(l)}\|\hat{R}^{(l-1)}\|_F, \tag{22}$$

where

$$\tilde{\omega}^{(l)} = \theta\omega^{(l)}\frac{1+\eta}{1-\eta}.$$

As we observe, for $\eta \in (0, (\sqrt{(\theta^2+1)^2+4\theta^2} - (\theta^2+1))/2\theta^2)$ we have $0 < (1+\eta)\eta\theta^2/(1-\eta) < 1$. So, for $\tilde{\omega} \in ((1+\eta)\eta\theta^2/(1-\eta), 1)$, we have $\tilde{\omega}^{(l)} \le \tilde{\omega}$ if

$$k_l \ge \frac{\ln((\tilde{\omega}(1-\eta)-\eta\theta^2(1+\eta))/2\theta^2(1+\eta)^2)}{\ln((\theta-1)/(\theta+1))}, \quad l = 1, 2, 3, \ldots,$$

Under this restriction, from Equation (22), we have

$$\begin{aligned}\|\hat{R}^{(l)}\|_F &\le \tilde{\omega}^{(l)}\|\hat{R}^{(l-1)}\|_F\\ &\le \Pi_{k=0}^{l}\tilde{\omega}^{(k)}\|\hat{R}^{(0)}\|_F\\ &\le \tilde{\omega}^{l+1}\|\hat{R}^{(0)}\|_F.\end{aligned}$$

Therefore, the residual sequence $\{R^{(l)}\}_{l=0}^{\infty}$ converges to zero matrix.

The above analysis is summarized in the following theorem.

Theorem 3.7: *Consider the Sylvester equation (1). Let $H_A(\alpha) = H_A - \alpha I_n$, $S_A(\alpha) = S_A + \alpha I_n$, $H_B(\alpha) = H_B - \alpha I_m$, $S_B(\alpha) = S_B + \alpha I_m$ and H_A, S_A, H_B, S_B are the Hermitian and skew-Hermitian parts of coefficient matrices A and B, respectively. Suppose that $\eta < 1$, and the NS-CGNR method is started from an initial guess $X^{(0)} \in \mathbb{R}^{n\times m}$, and produces an iterative sequence $\{X^{(l)}\}_{l=0}^{\infty}$, where $X^{(l)} \in \mathbb{R}^{n\times m}$ is the* l*th approximation to the solution $X^* \in \mathbb{R}^{n\times m}$ of the Sylvester equations* (1)*, by solving the Sylvester equation* (12) *with k_l steps of the Sylvester CGNR iterations. If $\hat{E}^{(l)}$ and $\hat{R}^{(l)}$ are as in* (18) *and* (21)*, respectively, then*

(a) $\|\hat{E}^{(l)}\|_F \le \omega^{(l)}\|\hat{E}^{(l-1)}\|_F$, $l = 1, 2, 3, \ldots,$
(b) $\|\hat{R}^{(l)}\|_F \le \tilde{\omega}^{(l)}\|\hat{R}^{(l-1)}\|_F$, $l = 1, 2, 3, \ldots,$

where

$$\omega^{(l)} = \left(2\left(\frac{\theta-1}{\theta+1}\right)^{k_l}(1+\eta)+\eta\right)\theta, \quad \tilde{\omega}^{(l)} = \theta\omega^{(l)}\frac{1+\eta}{1-\eta}, \quad l = 1, 2, 3, \ldots,$$

$$\theta = \frac{\max|\lambda(S_A(\alpha))| + \max|\lambda(S_B(\alpha))|}{\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|} \quad \text{and} \quad \eta = \theta^4\frac{\lambda_{\max}(H_A(\alpha)) + \lambda_{\max}(H_B(\alpha))}{\min|\lambda(S_A(\alpha))| + \min|\lambda(S_B(\alpha))|}.$$

Moreover, if $\eta \in (0, 1/\theta)$ then for some $\omega \in (\eta\theta, 1)$, and

$$k_l \ge \frac{\ln((\omega - \eta\theta)/(2\theta(1+\eta)))}{\ln((\theta-1)/(\theta+1))}, \quad l = 1, 2, 3, \ldots,$$

we have $\omega^{(l)} \le \omega (l = 1, 2, 3, \ldots)$, and the sequence $\{X^{(l)}\}_{l=0}^{\infty}$ converges to the solution X^* of the Sylvester equation (1)*. For $\eta \in (0, (\sqrt{(\theta^2+1)^2 + 4\theta^2} - (\theta^2+1))/2\theta^2)$ and some $\tilde{\omega} \in ((1+\eta)\eta\theta^2/(1-\eta), 1)$, and*

$$k_l \ge \frac{\ln((\tilde{\omega}(1-\eta) - \theta^2\eta(1+\eta))/(2\theta^2(1+\eta)^2))}{\ln((\theta-1)/(\theta+1))}, \quad l = 1, 2, 3, \ldots,$$

we have $\tilde{\omega}^{(l)} \le \tilde{\omega}(l = 1, 2, 3, \ldots)$, and the residual sequence $\{R^{(l)}\}_{l=0}^{\infty}$ converges to zero matrix.

3.3. Using the quasi-Hermitian splitting as a preconditioner

From the fact that any matrix splitting can naturally induce a splitting preconditioner for the Krylov subspace methods (see [6]), in Section 4, by numerical computation, we show that the quasi-Hermitian splitting (5) can be used as a splitting preconditioner and induce accurate, robust and effective preconditioned Krylov subspace iteration methods for solving the continuous Sylvester equation.

4. Numerical experiments

All numerical experiments presented in this section were computed in double precision with a number of MATLAB codes. All iterations are started from the zero matrix for initial $X^{(0)}$ and terminated when the current iterate satisfies $\|R^{(k)}\|_F \leq 10^{-8}\|R^{(0)}\|_F$, where $R^{(k)} = C - AX^{(k)} - X^{(k)}B$ is the residual of the kth iterate. In all problems, the right-hand side matrix C is chosen such that the solution of the Sylvester equation be a matrix with all entries equal to one. Also, we use the tolerance $\eta = 0.01$ for inner iterations in the NS-CGNR method. Inspired by the work of Golub and Vandrestraeten [16], we choose the parameter α similar to Equation (8). For each experiment we report the number of outer iterations denoted by 'out-itr', the number of total iterations denoted by 'tot-itr' and the average number of inner iterations denoted by 'av-initr'. The CPU time of each method is reported, too. Notation '†' in tables shows that the corresponding method is divergent. We compare the NS-CGNR method with some iterative methods. The iterative methods which used in this section are presented in Table 1.

Example 4.1: Consider the matrices

$$A = B = M + 2rN + \frac{100}{(n+1)^2}I,$$

where $M = \mathrm{tridiag}(-1, 2, -1)$, $N = \mathrm{tridiag}(0.5, 0, -0.5)$ and $n = 128$ [2]. When the preconditioned Krylov subspace iteration methods are used to solve the systems of linear equations resulting from the finite difference or the Sinc-Galerkin discretization of various differential equations and boundary value problems, this class of problems may arise [2].

We apply the methods to this problem for two values $r = 0.01$ and $r = 1$, and the results are given in Tables 2 and 3, respectively.

The case $r = 0.01$, is a problem of strong Hermitian part [24,31], but in the NS-CGNR method the skew-Hermitian part is emphasized. Therefore, we anticipate the NS-CGNR method has no efficiency or is divergent for this problem. The results in Table 2 have confirmed this idea. In this case the NSCG method is more effective versus the other methods.

The case $r = 1$, is a problem of strong skew-Hermitian part [24,31]. For this problem the NSCG method is divergent. Because, in the NSCG method the Hermitian part of coefficient matrices is

Table 1. Description of the used iterative methods.

Method	Description
NS-CGNR	New method described in Section 3
NSCG	Nested splitting conjugate gradient method described in [23]
HSS	Hermitian and skew-Hermitian splitting method described in [2]
IHSS	Inexact HSS method described in [2]
BiCGSTAB	BiCGSTAB method for the Sylvester equation, see [14,32]
PreBiCGSTAB	BiCGSTAB preconditioned by NS-CGNR
GMRES	GMRES(m) method for the Sylvester equation with $m = 10$, see [19,26]
FGMRES	FGMRES(m) preconditioned by NS-CGNR with $m = 10$

Table 2. Results for Example 4.1 with $r = 0.01$.

Method	CPU time	out-itr	tot-itr	av-initr
NS-CGNR	63.898	9430	9430	1
NSCG	1.045	7	452	64.57
HSS	54.910	192	5415	28.20
IHSS	14.570	209	1244	5.95
BiCGSTAB	3.447	146	–	–
GMRES	4.196	52	520	10

Table 3. Results for Example 4.1 with $r = 1$.

Method	CPU time	out-itr	tot-itr	av-initr
NS-CGNR	4.492	724	909	1.25
NSCG	†	†	†	†
HSS	721.832	172	43186	251.08
IHSS	173.043	179	10127	56.57
BiCGSTAB	11.494	501	–	–
PreBiCGSTAB	5.803	225	–	–
GMRES	3.104	39	390	10
FGMRES	4.041	29	290	10

emphasized [22,23]. Respect to the results presented in Table 3, we can observe that the NS-CGNR method is more effective than the other methods except the GMRES method. Moreover, the use of the quasi-Hermitian splitting as a preconditioner can improve the efficiency of the BiCGSTAB method, and reduce the number of iterations in the GMRES method. Figure 1 represents the effect of using the quasi-Hermitian splitting as a preconditioner in this problem.

Example 4.2: Consider symmetric matrix $H_n = n^2\text{pentadiag}(-1,-1,4,-1,-1) \in \mathbb{R}^{n\times n}$. Suppose that the skew-symmetric matrix $S_n \in \mathbb{R}^{n\times n}$ is a block diagonal matrix where every block is given by

$$S_{ii} = 2n \times \text{tridiag}(-1,0,1) \quad \text{for } i = 1,2,\ldots,n.$$

Now, let $A = H_n + 10^5 S_n$ and $B = H_m - 10^3 S_m$ with $n = 512$ and $m = 8$.

The results of this problem are given in Table 4.

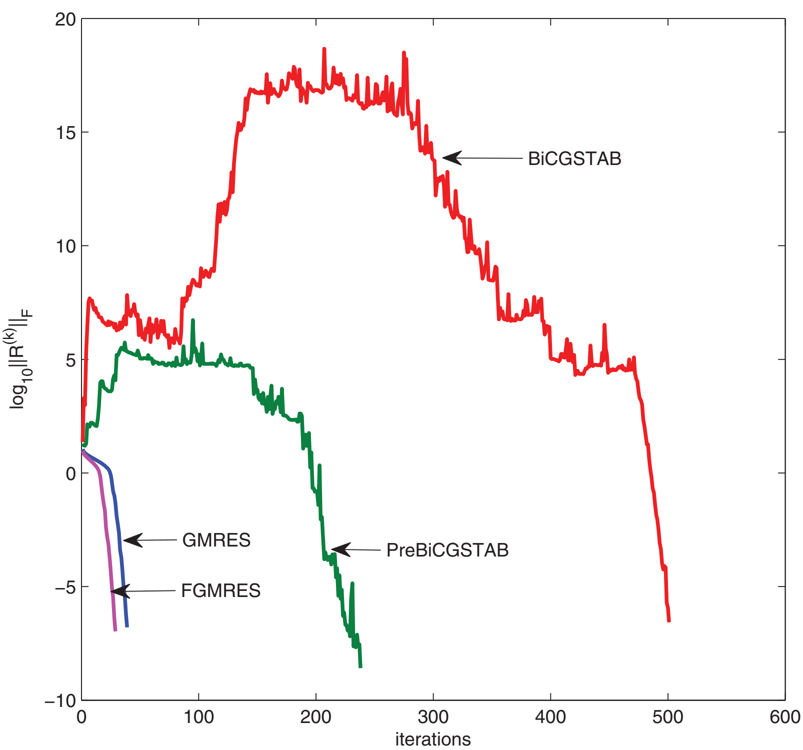

Figure 1. Effect of using the quasi-Hermitian splitting as a preconditioner in problem 4.1 with $r = 1$.

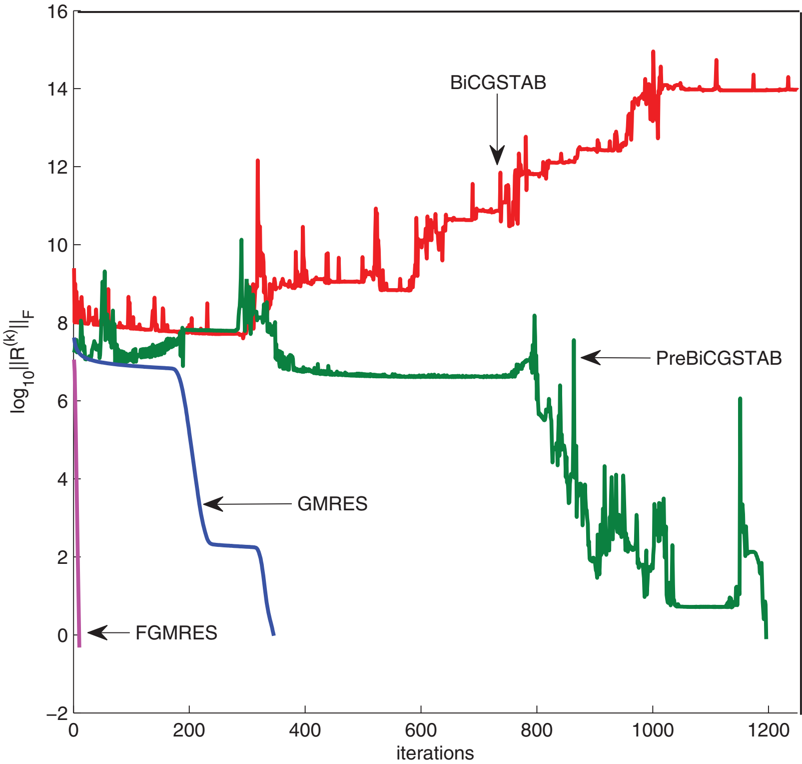

Figure 2. Effect of using the quasi-Hermitian splitting as a preconditioner in problem 4.2.

Table 4. Results for Example 4.2.

Method	CPU time	out-itr	tot-itr	av-initr
NS-CGNR	1.965	59	4067	68.93
NSCG	†	†	†	†
HSS	†	†	†	†
IHSS	†	†	†	†
BiCGSTAB	†	†	†	†
PreBiCGSTAB	52.743	1196	–	–
GMRES	2.558	346	3458	9.99
FGMRES	4.231	10	100	10

For this test problem, Table 4 shows that the NSCG, the HSS, the IHSS and the BiCGSTAB methods were diverging. According to the results presented in the Table 4, we can assert that the NS-CGNR is the most effective method among the considered methods. Especially, predominance is obvious in term of CPU time. Moreover, the use of the quasi-Hermitian splitting as a preconditioner can improve the efficiency of the BiCGSTAB method, and reduce the number of iterations in the GMRES method. We can observe desirableness and efficiency of using the quasi-HSS as a preconditioner splitting for the BiCGSTAB and the GMRES methods for this problem in Figure 3.

Example 4.3: Consider the matrices

$$A = B = H_n + 10^6 S_n,$$

where H_n and S_n are as in Example 4.2 and $n = 256$.

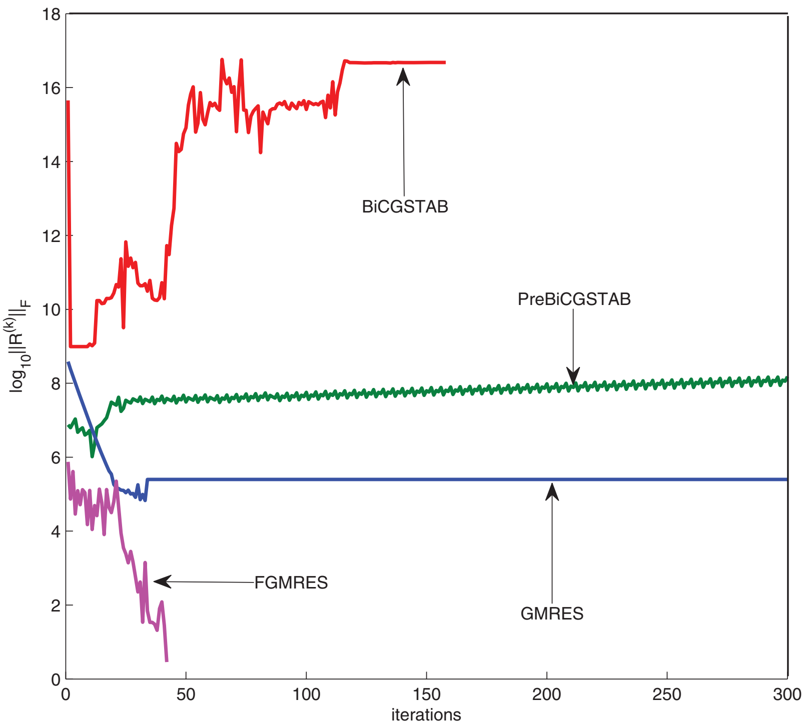

Figure 3. Effect of using the quasi-Hermitian splitting as a preconditioner in problem 4.3.

Table 5. Results for Example 4.3.

Method	CPU time	out-itr	tot-itr	av-initr
NS-CGNR	1468.343	1219	121664	99.80
NSCG	†	†	†	†
HSS	†	†	†	†
IHSS	†	†	†	†
BiCGSTAB	†	†	†	†
PreBiCGSTAB	†	†	†	†
GMRES	†	†	†	†
FGMRES	756.745	42	392	9.33

For this problem we have $\|H\|_2 = 29.7446$ and $\|S\|_2 = 1.4745e+07$, i.e., $\|S\|_2 >> \|H\|_2$. Therefore, this is a problem with strong skew-Hermitian part [22,23] and is a more challenging problem. The results of this problem are given in Table 5.

For this test problem, Table 5 shows that all methods, except the NS-CGNR method, were diverging. According to the results presented in the Table 5, we can assert that the use of the quasi-Hermitian splitting as a preconditioner cannot improve the efficiency of the BiCGSTAB method, but improve the efficiency of the GMRES method. We can observe desirableness and efficiency of using the quasi-HSS as a preconditioner splitting for the GMRES method for this problem in Figure 3.

5. Conclusion

In this paper, we have proposed an efficient iterative method for solving the continuous Sylvester equation $AX + XB = C$, by emphasizing the role of skew-Hermitian part of the coefficient matrices. This method which named NS-CGNR method, is suitable for problems with strong skew-Hermitian part. The NS-CGNR employs a CGNR-like method as inner iteration to approximate each outer iterate, while each outer iteration is induced by a convergent splitting of the coefficient matrices. Via this method, a Sylvester equation with coefficient matrices S_A and S_B (which are the skew-Hermitian part of A and B, respectively) is solved iteratively by a CGNR-like method.

We have compared the NS-CGNR method with some iterative methods such as the NSCG method, the HSS and the IHSS method, the BiCGSTAB method, and the GMRES method for some problems. We have observed that, for the problems with strong skew-Hermitian part, the NS-CGNR is superior the other iterative methods. However, we observe that the NS-CGNR method is not suitable for problems with the strong Hermitian part. In addition, numerical computations showed that the quasi-Hermitian splitting can induce the accurate, robust and effective preconditioned Krylov subspace method.

Acknowledgments

The authors are very much indebted to the referees for their constructive comments and suggestions which greatly improved the original manuscript of this paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] O. Axelsson, Z.-Z. Bai, and S.-X. Qiu, *A class of nested iterative schemes for linear systems with a coefficient matrix with a dominant positive definite symmetric part*, Numer. Algorithms 35 (2004), pp. 351–372.
[2] Z.-Z. Bai, *On Hermitian and skew-Hermitian splitting iteration methods for continuous Sylvester equations*, J. Comput. Math. 29 (2011), pp. 185–198.
[3] Z.-Z. Bai, G. H. Golub, and M. K. Ng, *Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems*, SIAM J. Matrix Anal. Appl. 24 (2003), pp. 603–626.
[4] Z.-Z. Bai, *Splitting iteration methods for non-Hermitian positive definite systems of linear equations*, Hokkaido Math. J. 36 (2007), pp. 801–814.
[5] Z.-Z. Bai, G. H. Golub, and M. K. Ng, *On inexact Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems*, Linear Algebr Appl. 428 (2008), pp. 413–440.
[6] Z.-Z. Bai, J.-F. Yin, and Y.-F. Su, *A shift-splitting preconditioner for non-Hermitian positive definite matrices*, J. Comput. Math. 24 (2006), pp. 539–552.
[7] R. H. Bartels and G. W. Stewart, *Algorithm 432: Solution of the matrix equation* $AX+XB=C$, Circ. Syst. Signal Proc. 13 (1994), pp. 820–826.
[8] P. Benner, R.-C. Li, and N. Truhar, *On the ADI method for Sylvester equations*, J. Comput. Appl. Math. 233 (2009), pp. 1035–1045.
[9] A. Bouhamidi and K. Jbilou, *Sylvester Tikhonov-regularization methods in image restoration*, J. Comput. Appl. Math. 206 (2007), pp. 86–98.
[10] D. Calvetti and L. Reichel, *Application of ADI iterative methods to the restoration of noisy images*, SIAM J. Matrix Anal. Appl. 17 (1996), pp. 165–186.
[11] B. Datta, *Numerical Methods for Linear Control Systems*, Elsevier Academic Press, New York, 2004.
[12] D.J. Evans and C.R. Wan, *A preconditioned conjugate gradient method for* $AX + XB = C$, Int. J. Comput. Math. 49 (1993), pp. 207–219.
[13] A. El Guennouni, K. Jbilou, and J. Riquet, *Block Krylov subspace methods for solving large Sylvester equation*, Numer. Algorithms 29 (2002), pp. 75–96.
[14] A. El Guennouni, K. Jbilou, and H. Sadok, *A block version of BiCGSTAB for linear systems with multiple right-hand sides*, Electron. Trans. Numer. Anal. 16 (2004), pp. 243–256.
[15] G.H. Golub, S. Nash, and C. Van Loan, *A Hessenberg–Schur method for the problem* $AX+XB=C$, IEEE Trans. Contr. AC-24 (1979), pp. 909–913.
[16] G.H. Golub and D. Vandrestraeten, *On the preconditioning of matrices with skew-symmetric splittings*, Numer. Algorithms 25 (2000), pp. 223–239.
[17] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[18] D.Y. Hu and L. Reichel, *Krylov-subspace methods for the Sylvester equation*, Linear Algebr. Appl. 172 (1992), pp. 283–313.
[19] K. Jbilou, A. Messaoudi, and H. Sadok, *Global FOM and GMRES algorithms for matrix equations*, Appl. Numer. Math. 31 (1999), pp. 97–109.

[20] Y.-F. Ke and C.-F. Ma, *A preconditioned nested splitting conjugate gradient iterative method for the large sparse generalied Sylvester equation*, Comput. Math. Appl. 68 (2014), pp. 1409–1420.
[21] C.T. Kelley, *Iterative Methods for Linear and Nonlinear Equations, vol. 16*, Frontiers in Applied Mathematics, SIAM, Philadelphia, 1995.
[22] M. Khorsand Zak and F. Toutounian, *Nested splitting conjugate gradient method for matrix equation $AXB = C$ and preconditioning*, Comput. Math. Appl. 66 (2013), pp. 269–278.
[23] M. Khorsand Zak and F. Toutounian, *Nested splitting CG-like iterative method for solving the continuous Sylvester equation and preconditioning*, Adv. Comput. Math. 40 (2014), pp. 865–880.
[24] L.A. Krukier, T.S. Martynova, and Z.-Z. Bai, *Product-type skew-Hermitian triangular splitting iteration methods for strongly non-Hermitian positive definite linear systems*, J. Comput. Appl. Math. 232 (2009), pp. 3–16.
[25] H. Lütkepohl, *Handbook of Matrices*, John Wiley & Sons Press, England, 1996.
[26] D. K. Salkuyeh and F. Toutounian, *New approaches for solving large Sylvester equations*, Appl. Math. Comput. 173 (2006), pp. 9–18.
[27] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
[28] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, 3rd ed., Springer-Verlag, New York, 2002.
[29] C.-L. Wang and Z.-Z. Bai, *Sufficient conditions for the convergent splittings of non-Hermitian positive definite matrices*, Linear Algebr. Appl. 330 (2001), pp. 215–218.
[30] L. Wang and Z.-Z. Bai, *Convergence conditions for splitting iteration methods for non-Hermitian linear systems*, Linear Algebr. Appl. 428 (2008), pp. 453–468.
[31] L. Wang and Z.-Z. Bai, *Skew-Hermitian triangular splitting iteration methods for non-Hermitian positive definite linear systems of strong skew-Hermitian parts*, BIT Numer. Math. 44 (2004), pp. 363–386.
[32] M. Wang and Y. Feng, *An iterative algorithm for solving a class of matrix equations*, J. Control Theory Appl. 7 (2009), pp. 68–72.

Appendix

To prove Theorem 2.1, we need the following lemmas.

Lemma A.1 ([28]): *If* F *is an $n \times n$ matrix with $\|F\| < 1$, then $(I + F)^{-1}$ exists and satisfies*

$$\|(I + F)^{-1}\| \leq \frac{1}{1 - \|F\|}.$$

Lemma A.2 ([27]): *Let $\mathcal{A} \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix, and assume that the system of linear equations* (2) *is solved by the conjugate gradient method. If $x^{(0)} \in \mathbb{R}^n$ is the starting vector, $x^{(k)} \in \mathbb{R}^n$ the* k*th iterate, and $x^* \in \mathbb{R}^n$ the exact solution of the linear system of equations* (2), *then*

$$\|x^{(k)} - x^*\|_{\mathcal{A}} \leq 2\left(\frac{\sqrt{\kappa(\mathcal{A})} - 1}{\sqrt{\kappa(\mathcal{A})} + 1}\right)^k \|x^{(0)} - x^*\|_{\mathcal{A}}.$$

Evidently, by making use of Lemma .2, we can obtain the following corollary:

Corollary A.3: *Let $\mathcal{A} \in \mathbb{R}^{n\times n}$, and assume that the system of linear equations* (2) *is solved by the CGNR method. If $x^{(0)} \in \mathbb{R}^n$ is the starting vector, $x^{(k)} \in \mathbb{R}^n$ the* k*th iterate, and $x^* \in \mathbb{R}^n$ the exact solution of the system of linear equations* (2), *then*

$$\|x^{(k)} - x^*\|_{\mathcal{A}^{\mathrm{T}}\mathcal{A}} \leq 2\left(\frac{\kappa(\mathcal{A}) - 1}{\kappa(\mathcal{A}) + 1}\right)^k \|x^{(0)} - x^*\|_{\mathcal{A}^{\mathrm{T}}\mathcal{A}}.$$

Now, we can prove the Theorem 2.1 as follows:

Proof of Theorem 2.1: Let $x^{(*,l)}$ be the exact solution of the system of linear equations (6). Then it satisfies

$$x^{(*,l)} = \mathcal{S}_\alpha^{-1}c - \mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha x^{(l-1)}.$$

On the other hand, since x^* is the exact solution of the system of linear equations (2), it obeys

$$x^* = \mathcal{S}_\alpha^{-1}c - \mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha x^*.$$

Let $\mu(\mathcal{S}_\alpha) = (\kappa(\mathcal{S}_\alpha) - 1)/(\kappa(\mathcal{S}_\alpha) + 1)$. Then according to Corollary .3, we have

$$\begin{aligned}
\|x^{(l)} - x^{(*,l)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} &\leq 2\mu(\mathcal{S}_\alpha)^{k_l}\|x^{(l-1)} - x^{(*,l)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&= 2\mu(\mathcal{S}_\alpha)^{k_l}\|x^{(l-1)} - (\mathcal{S}_\alpha^{-1}c - \mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha x^{(l-1)})\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&= 2\mu(\mathcal{S}_\alpha)^{k_l}\|(I + \mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha)x^{(l-1)} - (I + \mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha)x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&= 2\mu(\mathcal{S}_\alpha)^{k_l}\|(I + \mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha)(x^{(l-1)} - x^*)\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&\leq 2\mu(\mathcal{S}_\alpha)^{k_l}(1 + \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha})\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&= 2\mu(\mathcal{S}_\alpha)^{k_l}(1 + \varrho)\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}.
\end{aligned}$$

Furthermore, we can obtain

$$\begin{aligned}
\|x^{(l)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} &= \|(x^{(l)} - x^{(*,l)}) - (x^{(*,l)} - x^*)\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&\leq \|x^{(l)} - x^{(*,l)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} + \|x^{(*,l)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&\leq 2\mu(\mathcal{S}_\alpha)^{k_l}(1 + \varrho)\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} + \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha(x^{(l-1)} - x^*)\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&\leq 2\mu(\mathcal{S}_\alpha)^{k_l}(1 + \varrho)\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} + \varrho\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&= (2\mu(\mathcal{S}_\alpha)^{k_l}(1 + \varrho) + \varrho)\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&= \gamma^{(l)}\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}.
\end{aligned}$$

This proves the validity of (a). We now turn to the proof of (b). Since

$$\|c - \mathcal{A}x^{(l)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} = \|\mathcal{A}(x^{(l)} - x^*)\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \leq \|\mathcal{A}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}\|x^{(l)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha},$$

making use of (a) we have

$$\begin{aligned}
\|c - \mathcal{A}x^{(l)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} &\leq \gamma^{(l)}\|\mathcal{A}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&\leq \gamma^{(l)}\|\mathcal{A}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}\|\mathcal{A}^{-1}(c - \mathcal{A}x^{(l-1)})\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&\leq \gamma^{(l)}\|\mathcal{A}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}\|\mathcal{A}^{-1}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}\|c - \mathcal{A}x^{(l-1)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}.
\end{aligned}$$

We easily obtain

$$\begin{aligned}
\|\mathcal{A}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} &= \|\mathcal{H}_\alpha + \mathcal{S}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&= \|\mathcal{S}_\alpha(\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha + I)\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&\leq \|\mathcal{S}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}(1 + \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}),
\end{aligned}$$

and since $\|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} < 1$, by Lemma .1, we can obtain

$$\begin{aligned}
\|\mathcal{A}^{-1}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} &= \|(\mathcal{H}_\alpha + \mathcal{S}_\alpha)^{-1}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&= \|(\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha + I)^{-1}\mathcal{S}_\alpha^{-1}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&\leq \|\mathcal{S}_\alpha^{-1}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}\|(\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha + I)^{-1}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\
&\leq \frac{\|\mathcal{S}_\alpha^{-1}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}}{1 - \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}}.
\end{aligned}$$

On the other hand, we have $\|\mathcal{S}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} = \|\mathcal{S}_\alpha\|_2$ and $\|\mathcal{S}_\alpha^{-1}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} = \|\mathcal{S}_\alpha^{-1}\|_2$. Therefore, we have:

$$\|\mathcal{A}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \leq \|\mathcal{S}_\alpha\|_2(1 + \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}) \quad \text{and} \quad \|\mathcal{A}^{-1}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \leq \frac{\|\mathcal{S}_\alpha^{-1}\|_2}{1 - \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}}.$$

Therefore, it follows that

$$\begin{aligned}\|c - \mathcal{A}x^{(l)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} &\leq \gamma^{(l)} \frac{\|\mathcal{S}_\alpha\|_2(1 + \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha})\|\mathcal{S}_\alpha^{-1}\|_2}{1 - \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}} \|c - \mathcal{A}x^{(l-1)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\ &= \gamma^{(l)}\kappa(\mathcal{S}_\alpha)\frac{1 + \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}}{1 - \|\mathcal{S}_\alpha^{-1}\mathcal{H}_\alpha\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}} \|c - \mathcal{A}x^{(l-1)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\ &= \tilde{\gamma}^{(l)}\|c - \mathcal{A}x^{(l-1)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha}\end{aligned}$$

This shows the validity of (b).

It is obvious that, for $\gamma \in (\varrho, \varrho_1)$ with $\varrho_1 = \min\{1, 2 + 3\varrho\}$, $\gamma^{(l)} \leq \gamma$ $(l = 1, 2, \ldots)$ holds under condition

$$k_l \geq \frac{\ln((\gamma - \varrho)/2(1 + \varrho))}{\ln((\kappa(\mathcal{S}_\alpha) - 1)/(\kappa(\mathcal{S}_\alpha) + 1))}, \quad l = 1, 2, \ldots,$$

and the estimates

$$\begin{aligned}\|x^{(l)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} &\leq \gamma^{(l)}\|x^{(l-1)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\ &\leq \Pi_{k=0}^{l}\gamma^{(k)}\|x^{(0)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\ &\leq \gamma^{l+1}\|x^{(0)} - x^*\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \to 0, \quad l \to \infty,\end{aligned}$$

hold in accordance with (a). Therefore, the sequence $\{x^{(l)}\}_{l=0}^{\infty}$ converges to the solution x^* of the system of linear equations (2).

In addition, for $\varrho \in (0, r)$, where r is the positive root of quadratic equation $\kappa(\mathcal{S}_\alpha)\varrho^2 + (\kappa(\mathcal{S}_\alpha) + 1)\varrho - 1 = 0$ and $0 < r < 1$, we have $0 < (1 + \varrho)\varrho\kappa(\mathcal{S}_\alpha)/(1 - \varrho) < 1$. So, for

$$\tilde{\gamma} \in \left(\frac{(1 + \varrho)\varrho}{(1 - \varrho)}\kappa(\mathcal{S}_\alpha), 1\right),$$

$\tilde{\gamma}^{(l)} \leq \tilde{\gamma}$ $(l = 1, 2, 3, \ldots)$ holds under condition

$$k_l \geq \frac{\ln(((1 - \varrho)\tilde{\gamma} - \varrho(1 + \varrho)\kappa(\mathcal{S}_\alpha))/2(1 + \varrho)^2\kappa(\mathcal{S}_\alpha))}{\ln((\kappa(\mathcal{S}_\alpha) - 1)/(\kappa(\mathcal{S}_\alpha) + 1))}, \quad l = 1, 2, 3, \ldots,$$

and the estimates

$$\begin{aligned}\|c - \mathcal{A}x^{(l)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} &\leq \tilde{\gamma}^{(l)}\|c - \mathcal{A}x^{(l-1)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\ &\leq \Pi_{k=0}^{l}\tilde{\gamma}^{(k)}\|c - \mathcal{A}x^{(0)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \\ &\leq \tilde{\gamma}^{l+1}\|c - \mathcal{A}x^{(0)}\|_{\mathcal{S}_\alpha^{\mathrm{T}}\mathcal{S}_\alpha} \to 0, \quad l \to \infty,\end{aligned}$$

hold in accordance with (b). ∎