Received: 13 April 2017 | Revised: 3 November 2017 | Accepted: 23 February 2018

DOI: 10.1002/nla.2176

SPECIAL ISSUE PAPER

WILEY

Krylov methods for low-rank commuting generalized Sylvester equations

Elias Jarlebring[1] | Giampaolo Mele[1] | Davide Palitta[2] | Emil Ringh[1]

[1]Department of Mathematics, KTH Royal Institute of Technology, Swedish e-science research center (SeRC), Lindstedtsvägen 25, SE-100 44 Stockholm, Sweden

[2]Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato, 5, I-40127, Bologna, Italy

Correspondence
Giampaolo Mele, Department of Mathematics, KTH Royal Institute of Technology, Swedish e-science research center (SeRC), Lindstedtsvägen 25, SE-100 44 Stockholm, Sweden.
Email: gmele@kth.se

Funding information
INdAM-GNCS; Swedish Research Council, Grant/Award Number: 621-2013-4640

JEL Classification: 39B42; 65F10; 58E25; 47A46; 65F30

Summary

We consider generalizations of the Sylvester matrix equation, consisting of the sum of a Sylvester operator and a linear operator Π with a particular structure. More precisely, the commutators of the matrix coefficients of the operator Π and the Sylvester operator coefficients are assumed to be matrices with low rank. We show (under certain additional conditions) low-rank approximability of this problem, that is, the solution to this matrix equation can be approximated with a low-rank matrix. Projection methods have successfully been used to solve other matrix equations with low-rank approximability. We propose a new projection method for this class of matrix equations. The choice of the subspace is a crucial ingredient for any projection method for matrix equations. Our method is based on an adaption and extension of the extended Krylov subspace method for Sylvester equations. A constructive choice of the starting vector/block is derived from the low-rank commutators. We illustrate the effectiveness of our method by solving large-scale matrix equations arising from applications in control theory and the discretization of PDEs. The advantages of our approach in comparison to other methods are also illustrated.

KEYWORDS

generalized Sylvester equation, iterative solvers, Krylov subspace, low-rank commutation, matrix equation, projection methods

1 | INTRODUCTION

Let $\mathcal{L} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ denote the *Sylvester operator* associated with the matrices $A, B \in \mathbb{R}^{n\times n}$, that is,

$$\mathcal{L}(X) := AX + XB^T, \tag{1}$$

and let $\Pi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ denote the matrix operator defined by

$$\Pi(X) := \sum_{i=1}^{m} N_i X M_i^T, \tag{2}$$

where $m \ll n$. The matrices A and B are assumed to be large, sparse, and nonsingular; the operator $\mathcal{L}$ is assumed to be invertible, that is, the spectra of A and $-B$ are disjoint (see section 7.2 of Simoncini[1]). Given $C_1, C_2 \in \mathbb{R}^{n\times r}$ with $r \ll n$, our paper concerns the problem of computing $X \in \mathbb{R}^{n\times n}$ such that

$$\mathcal{L}(X) + \Pi(X) = C_1 C_2^T. \tag{3}$$

This equation is sometimes (e.g., Benner et al.[2]) referred to as the *generalized Sylvester equation*.

Let $[A, B] := AB - BA$ denote the *commutator* of two matrices. The structure of the operator Π is assumed to be such that the commutator of the Sylvester coefficients and the coefficients defining the operator Π have low rank. In other words, we assume that there exist $U_i, \tilde{U}_i \in \mathbb{R}^{n\times s_i}$ and $Q_i, \tilde{Q}_i \in \mathbb{R}^{n\times t_i}$ such that $s_i, t_i \ll n$ and the commutators fulfill

$$[A, N_i] = AN_i - N_iA = U_i\tilde{U}_i^T, \tag{4a}$$

$$[B, M_i] = BM_i - M_iB = Q_i\tilde{Q}_i^T, \tag{4b}$$

for $i = 1, \dots, m$. The property (4), which we refer to as *low-rank commutation*, is a generalization of the concept of commuting matrices in this framework. The case of pure commutation, that is, when the right-hand side of (4) is zero, which occurs for instance when $N_i = f_i(A), M_i = g_i(B)$, where f_i, g_i are polynomials or analytic functions, is analyzed in Lancaster[3] and Benner et al.[4]

A recent and successful method class for matrix equations defined by large and sparse matrices is based on projection, typically called *projection methods*.[4–6] We propose a new projection method for (3) under the low-rank commutation assumption (4).

Projection methods are typically derived from an assumption on the decay of the singular values of the solution. More precisely, a necessary condition for the successful application of a projection method is low-rank approximability, that is, the solution can be approximated by a low-rank matrix. We characterize the low-rank approximability of the solution to (3) under the condition that the Sylvester operator $\mathcal{L}$ has a low-rank approximability property and that $\rho(\mathcal{L}^{-1}\Pi) < 1$. The low-rank approximability theory is presented in Section 2. The function $\rho(\cdot)$ denotes the (operator) spectral radius, that is, $\rho(\mathcal{L}) := \sup\{|\lambda| \,|\, \lambda \in \Lambda(\mathcal{L})\}$.

The choice of the subspace is an important ingredient in any projection method. We propose a particular choice of projection spaces by identifying certain features of the solution to (3) based on our characterization of low-rank approximability and the low-rank commutation properties (4). More precisely, we use an extended Krylov subspace with an appropriate choice of the starting block. We present and analyze an expansion of the framework of the extended Krylov subspace method for Sylvester equation (Krylov-plus-inverted Krylov or K-PIK)[5,7] to the generalized Sylvester equation (Section 3).

Linear matrix equations of the form (3) arise in different applications. For example, the *generalized Lyapunov equation*, which corresponds to the special case where $B = A$, $M_i = N_i$, and $C_1 = C_2$, arises in model order reduction of bilinear and stochastic systems (e.g., see Benner et al.,[2,4] Damm[8] and references therein). Many problems arising from the discretization of PDEs can be formulated as generalized Sylvester equations.[9–11] Low-rank approximability for matrix equations has been investigated in different settings: for Sylvester equations,[12–14] generalized Lyapunov equations with low-rank correction[4] and more generally for linear systems with tensor product structure.[14,15]

The so-called low-rank methods, which projection methods belong to, directly compute a low-rank approximation to the solution of (3). Many algorithms have been developed for the Sylvester equation: projection methods,[5,6] low-rank Alternating Direction Implicit (ADI),[16,17] sign function method,[18,19] Riemannian optimization methods,[20,21] and many more. See the thorough presentation by Simoncini.[1] For large-scale generalized Sylvester equations, fewer numerical methods are available in the literature. Moreover, they are often designed only for solving the generalized Lyapunov equation although they may be adapted to solve the generalized Sylvester equation. Benner et al.[4] proposed a bilinear ADI (BilADI) method that naturally extends the low-rank ADI algorithm for standard Lyapunov problems to generalized Lyapunov equations. A nonstationary iterative method is derived in Shank et al.,[22] and in Kressner et al.,[23] a greedy low-rank technique is presented. In principle, it is always possible to consider the $n^2 \times n^2$ linear system that stems from Equation (3) by Kronecker transformations. There are specific methods for solving linear systems with tensor product structure (see Kressner et al.,[20,23] Ballani et al.[24] and references therein). These problems can also be solved by employing one of the many methods for linear systems presented in the literature. In particular, matrix-equation-oriented versions of iterative methods for linear systems, together with preconditioning techniques, are present in the literature (e.g., see section 5 of Benner et al,[4] Kressner and Tobler,[15] Bouhamidi and Jbilou,[25] Li et al.[26]). To our knowledge, the low-rank commutativity properties (4) have not been considered in the literature in the context of methods for matrix equations.

The paper is structured as follows. In Section 2, we use a Neumann series (cf. Lancaster,[3] Flagg and Gugercin,[27] Richter et al.,[28] Zhang and Lam[29]). with hypothesis $\rho(\mathcal{L}^{-1}\Pi) < 1$ to characterize the low-rank approximability of the solution to (3). In Section 3, we further characterize the approximation properties of the solution to (3) by exploiting the low-rank commutation feature of the coefficients (4). We use this characterization in the derivation of an effective projection space. In Section 3.4, we present an efficient procedure for solving small-scale generalized Sylvester equations (3).

Numerical examples that illustrate the effectiveness of our strategy are reported in Section 4. Our conclusions are given in Section 5.

We use the following notation. The vectorization operator vec: $\mathbb{R}^{n\times n} \to \mathbb{R}^{n^2}$ is defined such that vec(A) is the vector obtained by stacking the columns of the matrix A on top of one another. We denote by $\|\cdot\|_F$ the Frobenius norm, whereas $\|\cdot\|$ is any submultiplicative matrix norm. For a generic linear and continuous operator $\mathcal{L} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, the induced norm is defined as $\|\mathcal{L}\| := \sup_{\|A\|=1}\|\mathcal{L}(A)\|$. The identity and the zero matrices are respectively denoted by I and O. We denote by e_i the ith vector of the canonical basis of $\mathbb{R}^n$, whereas $\otimes$ corresponds to the Kronecker product. The matrix obtained by stacking the matrices $A_1, \ldots, A_n$ next to each other is denoted by $(A_1, \ldots, A_n)$. In conclusion, Range(A) is the vector space generated by the columns of the matrix A, and span$(\mathcal{A})$ is the vector space generated by the vectors in the set $\mathcal{A}$.

2 | REPRESENTATION AND APPROXIMATION OF THE SOLUTION

2.1 | Representation as Neumann series expansion

The following theorem gives sufficient conditions for the existence of a representation of the solution to a generalized Sylvester equation (3) as a convergent series. This will be needed for the low-rank approximability characterization in the following section, as well as in the derivation of a method for small generalized Sylvester equations (further described in Section 3.4).

Theorem 1. (Solution as a Neumann series)
Let $\mathcal{L}, \Pi : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be linear operators such that $\mathcal{L}$ is invertible, $\rho(\mathcal{L}^{-1}\Pi) < 1$, and let $C \in \mathbb{R}^{n\times n}$. The unique solution of the equation $\mathcal{L}(X) + \Pi(X) = C$ can be represented as

$$X = \sum_{j=0}^{\infty} Y_j, \tag{5}$$

where

$$\begin{cases} Y_0 & := \mathcal{L}^{-1}(C), \\ Y_{j+1} & := -\mathcal{L}^{-1}\left(\Pi(Y_j)\right), \qquad j \geq 0. \end{cases} \tag{6}$$

Proof. By using the invertibility of $\mathcal{L}$, we have $X = (I + \mathcal{L}^{-1}\Pi)^{-1}\mathcal{L}^{-1}(C)$, and with the assumption $\rho(\mathcal{L}^{-1}\Pi) < 1$, we can express the operator $(I + \mathcal{L}^{-1}\Pi)^{-1}$ as a convergent Neumann series (for operators as, e.g., in example 4.5 of Kato[30]). In particular, we obtain

$$X = \sum_{j=0}^{\infty} (-1)^j \left(\mathcal{L}^{-1}\Pi\right)^j \mathcal{L}^{-1}(C).$$

The relation (5) follows by defining $Y_j := (-1)^j \left(\mathcal{L}^{-1}\Pi\right)^j \mathcal{L}^{-1}(C)$. By induction, it follows that the relations (6) are fulfilled. □

Remark 1. Theorem 1 can be used to construct an approximation to the solution of $\mathcal{L}(X) + \Pi(X) = C$ by truncating the series (5) analogous to the general form given in equation (4.23) of Kato.[30] In particular, let

$$X^{(\ell)} := \sum_{j=0}^{\ell} Y_j, \tag{7}$$

where Y_j are given by (6). The truncation error can be bounded as follows:

$$\left\|X - X^{(\ell)}\right\| \leq \left\|\mathcal{L}^{-1}(C)\right\| \frac{\rho(\mathcal{L}^{-1}\Pi)^{\ell+1}}{1 - \rho(\mathcal{L}^{-1}\Pi)}.$$

If $\mathcal{L}$ and Π are respectively the operators (1) and (2) that define the generalized Sylvester equation (3), then the truncated Neumann series (7) can be efficiently computed for small-scale problems. In particular, this approach can be used in the derivation of a numerical method for solving small-scale generalized Sylvester equations as illustrated in Section 3.4.

2.2 | Low-rank approximability

We now use the result in the previous section to show that the solution to (3) can be often approximated by a low-rank matrix. We base the reasoning on low-rank approximability properties of $\mathcal{L}$. Our result requires the explicit use of certain conditions on the spectrum of matrix coefficients of $\mathcal{L}$. Under these specific conditions, the solution to a Sylvester equation with low-rank right-hand side can be approximated by a low-rank matrix (see section 4.1 of Simoncini[1]). In this sense, we can extend several results concerning the low-rank approximability of the solution to Sylvester equations to the case of generalized Sylvester equations under the assumption $\rho(\mathcal{L}^{-1}\Pi) < 1$. More precisely, the truncated Neumann series (7) is obtained by summing the solutions to the Sylvester equations (6). Note that, under the low-rank approximability assumption of $\mathcal{L}$, the right-hand sides of the Sylvester equations (6) are low-rank matrices because we assume that C is a low-rank matrix and $m \ll n$. We formalize this argument and present a new characterization of the low-rank approximability of the solution to (3) by adapting one of the most commonly used low-rank approximability result for Sylvester equations.[14]

We now briefly recall some results presented by Grasedyck,[14] for our purposes. Suppose that the matrix coefficients representing $\mathcal{L}$ are such that $\lambda(A) \cup \lambda(B) \subset \mathbb{C}_-$. Let $M \in \mathbb{C}^{n\times n}$ be such that $\lambda(M) \subset \mathbb{C}_-$; then, its inverse can be expressed as $M^{-1} = \int_0^\infty \exp(tM)dt$. The integral can be approximated with the following quadrature formula:

$$M^{-1} = \int_0^\infty \exp(tM)dt \approx \sum_{j=-k}^{k} w_j \exp(t_j M), \tag{8}$$

where the weights w_j and nodes t_j are given in lemma 5 of Grasedyck.[14] More precisely, we have an explicit formula for the approximation error

$$\left\| \int_0^\infty \exp(tM)dt - \sum_{j=-k}^{k} w_j \exp(t_j M) \right\| \leq K e^{-\pi\sqrt{k}}, \tag{9}$$

where K is a constant that only depends on the spectrum of M. The solution to the Sylvester equation $\mathcal{L}(X) = C$ can be explicitly expressed as $\mathrm{vec}(X) = (I \otimes A + B \otimes I)^{-1}\mathrm{vec}(C)$. The solution to this linear system can be approximated by using (8) and by approximating the inverse of $I \otimes A + B \otimes I$. Let $\mathcal{L}_k^{-1} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be the linear operator such that $\mathcal{L}_k^{-1}(C)$ corresponds to the approximation (8). More precisely, the operator $\mathcal{L}_k^{-1}$ satisfies

$$\mathrm{vec}\left(\mathcal{L}_k^{-1}(C)\right) = \sum_{j=-k}^{k} w_j \left[\exp(t_j B) \otimes \exp(t_j A)\right] \mathrm{vec}(C).$$

By using the properties of the Kronecker product, it can be explicitly expressed as

$$\mathcal{L}_k^{-1}(C) = \sum_{j=-k}^{k} w_j \exp(t_j A) C \exp(t_j B^T). \tag{10}$$

In terms of operators, the error bound (9) is $\|\mathcal{L}^{-1} - \mathcal{L}_k^{-1}\| \leq K e^{-\pi\sqrt{k}}$. The result of the above discussion is summarized in the following remark, which directly follows from (10) or lemma 7 of Grasedyck[14] and lemma 2 of Benner et al.[4]

Remark 2. The solution to the Sylvester equation $\mathcal{L}(X) = C$ can be approximated by $\bar{X} = \mathcal{L}_k^{-1}(C)$, where $\|X - \bar{X}\| \leq \|C\| K e^{-\pi\sqrt{k}}$, $\mathrm{rank}(\bar{X}) \leq (2k+1)r$, K is a constant that depends on the spectrum of $\mathcal{L}$, and r is the rank of C.

The following theorem concerns the low-rank approximability of the solution to (3). More precisely, it provides a generalization of Remark 2 to the case of generalized Sylvester equations by using the Neumann series characterization in Theorem 1.

Theorem 2. (Low-rank approximability)
Let $\mathcal{L}$ be the Sylvester operator (1)*, Π the linear operator* (2)*, $C_1, C_2 \in \mathbb{R}^{n\times r}$, and k a positive integer. Let $X^{(\ell)}$ be the truncated Neumann series* (7)*. Then, there exists a matrix $\bar{X}^{(\ell)}$ such that*

$$\mathrm{rank}(\bar{X}^{(\ell)}) \leq (2k+1)r + \sum_{j=1}^{\ell} (2k+1)^{j+1} m^j r, \tag{11}$$

and

$$\left\|X^{(\ell)} - \bar{X}^{(\ell)}\right\| \leq \bar{K} e^{-\pi\sqrt{k}}, \tag{12}$$

where $\bar{K}$ is a constant that does not depend on k and only depends on $\mathcal{L}$ and ℓ.

Proof. Let $\mathcal{L}_k$ be the operator (10), and consider the sequence

$$\begin{cases} \bar{Y}_0 & := \mathcal{L}_k^{-1}\left(C_1 C_2^T\right), \\ \bar{Y}_{j+1} & := -\mathcal{L}_k^{-1}\left(\Pi(\bar{Y}_j)\right), \quad j \geq 0. \end{cases} \tag{13}$$

Define $\beta := \|\mathcal{L}^{-1}\Pi\|$ and $\beta_k := \|\mathcal{L}_k^{-1}\Pi\|$. By using Remark 2, we have

$$\begin{aligned} \left\|Y_{j+1} - \bar{Y}_{j+1}\right\| &\leq \left\|\mathcal{L}^{-1}\left(\Pi(Y_j)\right) - \mathcal{L}^{-1}\left(\Pi(\bar{Y}_j)\right)\right\| + \left\|\mathcal{L}^{-1}\left(\Pi(\bar{Y}_j)\right) - \mathcal{L}_k^{-1}\left(\Pi(\bar{Y}_j)\right)\right\| \\ &\leq \beta \left\|Y_j - \bar{Y}_j\right\| + K e^{-\pi\sqrt{k}} \|\Pi\| \left\|\bar{Y}_j\right\|. \end{aligned}$$

From the above expression, a simple recursive argument shows that

$$\left\|Y_{j+1} - \bar{Y}_{j+1}\right\| \leq \beta^{j+1} \left\|Y_0 - \bar{Y}_0\right\| + K e^{-\pi\sqrt{k}} \|\Pi\| \sum_{t=0}^{j} \beta^{j-t} \left\|\bar{Y}_t\right\|. \tag{14}$$

Using the submultiplicativity of the operator norm, it holds that $\|\bar{Y}_j\| = \|\mathcal{L}_k^{-1}(\Pi(\bar{Y}_{j-1}))\| \leq \beta_k \|\bar{Y}_{j-1}\|$. In particular, $\|\bar{Y}_j\| \leq \beta_k^j \|\mathcal{L}_k^{-1}\| \|C_1 C_2^T\|$, and therefore, by using Remark 2, from (14), it follows that

$$\left\|Y_{j+1} - \bar{Y}_{j+1}\right\| \leq \left\|C_1 C_2^T\right\| K \left[\beta^{j+1} + \|\Pi\| \left\|\mathcal{L}_k^{-1}\right\| \sum_{t=0}^{j} \beta^{j-t} \beta_k^t\right] e^{-\pi\sqrt{k}}. \tag{15}$$

Because $\mathcal{L}_k^{-1}$ converges to $\mathcal{L}^{-1}$, and by using the continuity of the operators, we have that $\|\mathcal{L}_k^{-1}\|$ and β_k are bounded by a constant independent of k. Therefore, from (15), it follows that there exists a constant K_{j+1} independent of k such that $\|Y_{j+1} - \bar{Y}_{j+1}\| \leq K_{j+1} e^{-\pi\sqrt{k}}$. The relation (12) follows by defining $\bar{X}^{(\ell)} := \sum_{j=0}^{\ell} \bar{Y}_j$ and observing

$$\left\|X^{(\ell)} - \bar{X}^{(\ell)}\right\| \leq \sum_{j=0}^{\ell} \left\|Y_j - \bar{Y}_j\right\| \leq e^{-\pi\sqrt{k}} \sum_{j=0}^{\ell} K_j = \bar{K} e^{-\pi\sqrt{k}},$$

where $\bar{K} := \sum_{j=0}^{\ell} K_j$. The upper bound (11) follows by Remark 2 iteratively applied to (13). □

We want to point out that although Theorem 2 provides an explicit procedure for constructing an approximation to the solution of (3), we later consider a different class of methods. Theorem 2 has only theoretical interest and is used to motivate the employment of low-rank methods in the solution of (3). Moreover, in the numerical simulations (Section 4), we have observed a decay in the singular values of the solution to (3) that it is faster than the one predicted by Theorem 2.

3 | STRUCTURE EXPLOITING KRYLOV METHODS

3.1 | Extended Krylov subspace method

In this section, we derive a method for (3) that belongs to the class called projection methods. We briefly summarize the adaption of the projection method approach in our setting. Projection methods for matrix equations are iterative algorithms based on constructing two sequences of nested subspaces of $\mathbb{R}^n$, that is, $\mathcal{K}_{k-1} \subset \mathcal{K}_k$ and $\mathcal{H}_{k-1} \subset \mathcal{H}_k$. Justified by the low-rank approximability of the solution, projection methods construct approximations (of the solution to (3)) of the form

$$X_k = \mathcal{V}_k Z_k \mathcal{W}_k^T, \tag{16}$$

where $\mathcal{V}_k$ and $\mathcal{W}_k$ are matrices with orthonormal columns representing respectively an orthonormal basis of $\mathcal{K}_k$ and $\mathcal{H}_k$. Note that low-rank approximability (in the sense illustrated in, e.g., Theorem 2) is a necessary condition for the success of an approximation of the type (16).

The matrix Z_k can be obtained by imposing the Galerkin orthogonality condition, namely, the residual

$$\mathcal{R}_k := AX_k + X_k B^T + \sum_{i=1}^{m} N_i X_k M_i^T - C_1 C_2^T \tag{17}$$

is such that $\mathcal{V}_k^T \mathcal{R}_k \mathcal{W}_k = 0$. This condition is equivalent to Z_k satisfying the following small and dense generalized Sylvester equation, usually referred to as the *projected problem*:

$$T_k Z_k + Z_k H_k^T + \sum_{i=1}^{m} G_{k,i} Z_k F_{k,i}^T = E_{k,1} E_{k,2}^T, \tag{18}$$

where

$$T_k := \mathcal{V}_k^T A \mathcal{V}_k, \quad H_k := \mathcal{W}_k^T B \mathcal{W}_k, \quad E_{k,1} = \mathcal{V}_k^T C_1, \quad E_{k,2} = \mathcal{W}_k^T C_2, \tag{19a}$$

$$G_{k,i} := \mathcal{V}_k^T N_i \mathcal{V}_k, \quad F_{k,i} := \mathcal{W}_k^T M_i \mathcal{W}_k, \quad i = 1, \ldots, m. \tag{19b}$$

The iterative procedure consists of expanding the spaces $\mathcal{K}_k$ and $\mathcal{H}_k$ until the norm of the residual matrix $\mathcal{R}_k$ (17) is sufficiently small.

A projection method is efficient only if the subspaces $\mathcal{K}_k$ and $\mathcal{H}_k$ are selected in a way that the projected matrix (16) is a good low-rank approximation to the solution without the dimensions of the spaces being large. One of the most popular choices of subspace is the extended Krylov subspace (although certainly not the only choice.[6,31]) Extended Krylov subspaces form the basis of the method called K-PIK.[5,7] For our purposes, it is natural to define extended Krylov subspaces with the notation of block Krylov subspaces (cf. section 6 of Gutknecht[32]). Given an invertible matrix $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{n\times r}$, the extended block Krylov subspace can be defined as the sum of two vector spaces, more precisely $E\mathcal{B}_k^{\square}(A, C) := \mathcal{B}_k^{\square}(A, C) + \mathcal{B}_k^{\square}(A^{-1}, A^{-1}C)$, where

$$\mathcal{B}_k^{\square}(A, C) := \operatorname{span}\left(\left\{p(A)Cw \mid \deg(p) \le k,\ w \in \mathbb{R}^{r\times r}\right\}\right) \subseteq \mathbb{C}^{n\times r}$$

denotes the block Krylov subspace, $p \in \mathbb{R}[x]$ is a polynomial, and $\deg(\cdot)$ is the degree function. This space can be also represented as

$$\mathcal{B}_k^{\square}(A, C) = \underbrace{\mathcal{B}_k(A, C) \times \cdots \times \mathcal{B}_k(A, C)}_{r \text{ times}},$$

where $\mathcal{B}_k(A, C) := \operatorname{span}(\{p(A)Cw \mid \deg(p) \le k,\ w \in \mathbb{R}^r\}) \subseteq \mathbb{C}^n$.

The extended Krylov subspace method is a projection method where $\mathcal{K}_k = E\mathcal{B}_k(A, \bar{C}_1)$, $\mathcal{H}_k = E\mathcal{B}_k(B, \bar{C}_2)$, and $\bar{C}_1$, $\bar{C}_2$ are called the starting blocks, which we will show how to select in our setting in Sections 3.2 and 3.3. The procedure is summarized in Algorithm 1 where the matrices L and R are the low-rank factors of (16), that is, they are such that $X_k = LR^T$. Notice that, in the case of generalized Lyapunov equations, the new blocks V_k and W_k are equal (hence also the basis matrices $\mathcal{V}_k$ and $\mathcal{W}_k$), and Algorithm 1 can be optimized accordingly.

Algorithm 1: Extended Krylov subspace method for generalized Sylvester equations.

input : Matrix coeff.: $A, B, N_1 \ldots, N_m, M_1, \ldots, M_m \in \mathbb{R}^{n\times n}$, $C_1, C_2 \in \mathbb{R}^{n\times r}$
Starting blocks: $\bar{C}_1 \in \mathbb{R}^{n\times \bar{r}_1}$ and $\bar{C}_2 \in \mathbb{R}^{n\times \bar{r}_2}$
Maximum number of iterations: d

output: Low-rank factors: L, R

Set $V_1 = \texttt{orth}((\bar{C}_1, A^{-1}\bar{C}_1))$, $W_2 = \texttt{orth}((\bar{C}_2, B^{-1}\bar{C}_2))$, $\mathcal{V}_0 = \mathcal{W}_0 = \emptyset$
for $k = 1, 2, \ldots, d$, **do**
  $\mathcal{V}_k = (\mathcal{V}_{k-1}, V_k)$ and $\mathcal{W}_k = (\mathcal{W}_{k-1}, W_k)$
  Compute $T_k, H_k, E_{k,1}, E_{k,2}, G_{k,i}, F_{k,i}$ according to (19a)–(19b)
  Solve the *projected problem* (18)
  Compute $\|\mathcal{R}_k\|_F$ according to (21)
  if $\|\mathcal{R}_k\|_F \le$ `tol`, **then**
    Break
  end
  Set $V_k^{(1)}$: first $\bar{r}_1$ columns of V_k; Set $V_k^{(2)}$: last $\bar{r}_1$ columns of V_k
  Set $W_k^{(1)}$: first $\bar{r}_2$ columns of W_k; Set $W_k^{(2)}$: last $\bar{r}_2$ columns of W_k
  $V'_{k+1} = (AV_k^{(1)}, A^{-1}V_k^{(2)})$ and $W'_{k+1} = (BW_k^{(1)}, B^{-1}W_k^{(2)})$
  $\widehat{V}_{k+1} \leftarrow$ block-orthogonalize V'_{k+1} w.r.t. $\mathcal{V}_k$
  $\widehat{W}_{k+1} \leftarrow$ block-orthogonalize W'_{k+1} w.r.t. $\mathcal{W}_k$
  $V_{k+1} = \texttt{orth}\widehat{V}_{k+1}$ and $W_{k+1} = \texttt{orth}\widehat{W}_{k+1}$
end
Compute the decomposition $Z_k = \widehat{L}\widehat{R}^T$
Return $L = \mathcal{V}_k\widehat{L}$ and $R = \mathcal{W}_k\widehat{R}$

Remark 3. The output of Algorithm 1 represents the factorization $X_k = LR^T$. Under the condition that $\|\mathcal{R}_k\|$ is small, X_k is an approximation of the solution to (3) such that $\text{rank}(X_k) \leq 2\min(\bar{r}_1, \bar{r}_2)k$, $\bar{r}_1 = \text{rank}(\bar{C}_1)$, $\bar{r}_2 = \text{rank}(\bar{C}_2)$. By construction, $\text{Range}(L) \subseteq E\mathcal{B}_k(A, \bar{C}_1)$ and $\text{Range}(R) \subseteq E\mathcal{B}_k(B, \bar{C}_2)$. As it has been shown (e.g., in Simoncini,[5] Breiten et al.[7]), an orthonormal basis of $E\mathcal{B}_k$ can computed by means of the block Arnoldi procedure. Moreover, for the case of the Sylvester equation, that is, $m = 0$, Algorithm 1 can be effectively applied with starting blocks $\bar{C}_1 = C_1$ and $\bar{C}_2 = C_2$.

A breakdown in Algorithm 1 may occur in two situations. During the generation of the basis of the extended Krylov subspaces, (numerical) loss of orthogonality may occur in Steps 9–11. This issue is present already for the Sylvester equation,[5,7] and we refer to Gutknecht[32] for a presentation of safeguard strategies that may mitigate the problem. We assume that the bases $\mathcal{V}_k$ and $\mathcal{W}_k$ have full rank. The other situation where a breakdown may occur is in Step 4. It may happen that the projected problem (18) is not solvable. For the Sylvester equation, the solvability of the projected problem is guaranteed by the condition that the fields of values of A and $-B$ are disjoint (see section 4.4.1 of Simoncini[1]). We extend this result, which provides a way to verify the applicability of the method (without carrying out the method). As illustrated in the following proposition, for the generalized Sylvester equation, we need an additional condition. Instead of using the field of values, it is natural to phrase this condition in terms of the ratio field of values (defined in, e.g., Einstein et al.[33]).

Proposition 1. *Consider the generalized Sylvester equation* (3)*, and assume that the fields of values of A and $-B$ are disjoint and that the ratio field of values of $\sum_{i=1}^m M_i \otimes N_i$ and $B \otimes I + I \otimes A$, that is,*

$$R\left(\sum_{i=1}^m M_i \otimes N_i, B \otimes I + I \otimes A\right) := \left\{ \left. \frac{y^H \left(\sum_{i=1}^m M_i \otimes N_i\right) y}{y^H (B \otimes I + I \otimes A)\, y} \right| y \in \mathbb{C}^{n^2} \backslash \{0\} \right\},$$

is strictly contained in the open unit disk. Then, the projected problem (18) *has a unique solution.*

Proof. Let $\mathcal{L}_{proj}(Z) := T_k Z + Z H_k^T$ and $\Pi_{proj}(Z) := \sum_{i=1}^m G_{k,i} Z F_{k,i}^T$. The projected problem (18) is equivalently written as $\mathcal{L}_{proj}(Z_k) + \Pi_{proj}(Z_k) = E_{k,1} E_{k,2}^T$. Because A and $-B$ have disjoint fields of values, $\mathcal{L}_{proj}$ is invertible (see section 4.4.1 of Simoncini[1]). From Theorem 1, we know that there exists a unique solution Z_k to (18) if $\rho(\mathcal{L}_{proj}^{-1}\Pi_{proj}) < 1$. This condition is equivalent to $|\lambda| < 1$, where $(\lambda, v) \in \mathbb{C} \times \mathbb{C}^{(kr)^2} \backslash \{0\}$ is any eigenpair of the following generalized eigenvalue problem:

$$\left(\sum_{i=1}^m F_{k,i} \otimes G_{k,i}\right) v = \lambda (H_k \otimes I + I \otimes T_k) v. \tag{20}$$

Using the properties of the Kronecker product, Equation (20) can be written as

$$\sum_{i=1}^m \left(W_k^T \otimes V_k^T\right)(M_i \otimes N_i)(W_k \otimes V_k)\, v = \lambda \left(W_k^T \otimes V_k^T\right)(B \otimes I + I \otimes A)(W_k \otimes V_k)\, v.$$

By multiplying the above equation from the left with v^H, we have that

$$|\lambda| = \left| \frac{x^H \left(\sum_{i=1}^m M_i \otimes N_i\right) x}{x^H (B \otimes I + I \otimes A)\, x} \right|, \quad x := (W_k \otimes V_k)\, v.$$

By using that $R\left(\sum_{i=1}^m M_i \otimes N_i, B \otimes I + I \otimes A\right)$ is strictly contained in the unit circle, we conclude that $|\lambda| < 1$. □

Observation 1. The computation of the matrices T_k, H_k (Step 3) and the orthogonalization of the new blocks V_{k+1}, W_{k+1} (Steps 9–11) can be efficiently performed as in section 3 of Simoncini,[5] where a modified Gram–Schmidt method is employed in the orthogonalization. The matrices $G_{k,i}$ and $F_{k,i}$ (Step 3) can be computed by extending the matrices $G_{k-1,i}$ and $F_{k-1,i}$ with a block-column and a block-row. Moreover, the matrix X_k is never explicitly formed. In particular, the Frobenius norm of the residual (17) can be computed as

$$\|\mathcal{R}_k\|_F^2 = \left\|\tau_{k+1}(e_k \otimes I_{2r})^T Z_k\right\|_F^2 + \left\|Z_k (e_k \otimes I_{2r})^T h_{k+1}^T\right\|_F^2. \tag{21}$$

This follows by replacing in (17) the following Arnoldi-like relations (see eq. 4 of Simoncini et al.[34]):

$$A\mathcal{V}_k = \mathcal{V}_k T_k + V_{k+1}\tau_{k+1}(e_k \otimes I_{2r})^T, \quad B\mathcal{W}_k = \mathcal{W}_k H_k + W_{k+1} h_{k+1}(e_k \otimes I_{2r})^T.$$

Remark 4. Algorithm 1 is a block method in the sense that it uses a block Arnoldi procedure to generate the basis matrices $\mathcal{V}_k$ and $\mathcal{W}_k$. The basis generated by Algorithm 1 can hence be interpreted in a block sense and be related to the block Krylov spaces $E\mathcal{B}_k^{\square}(A, \bar{C}_1)$ and $E\mathcal{B}_k^{\square}(B, \bar{C}_2)$ (cf. section 2 of Frommer et al.[35]). However, in the framework of projection methods for matrix equations, the spans of the columns of $\mathcal{V}_k$ and $\mathcal{W}_k$ are often considered as the projection spaces. Moreover, the columns of $\mathcal{V}_k$ and $\mathcal{W}_k$ are respectively a basis for $E\mathcal{B}_k(A, \bar{C}_1)$ and $E\mathcal{B}_k(B, \bar{C}_2)$. In particular, each column of L and R is respectively in the space $E\mathcal{B}_k(A, \bar{C}_1)$ and $E\mathcal{B}_k(B, \bar{C}_2)$. This is equivalently expressed as $\text{Range}(L) \subseteq E\mathcal{B}_k(A, \bar{C}_1)$ and $\text{Range}(R) \subseteq E\mathcal{B}_k(B, \bar{C}_2)$. Therefore, in the following analysis, we derive and use properties of the spaces $E\mathcal{B}_k(A, \bar{C}_1)$ and $E\mathcal{B}_k(B, \bar{C}_2)$.

3.2 | Krylov subspace and low-rank commuting matrices

Algorithm 1 is efficient only if the starting blocks $\bar{C}_1$ and $\bar{C}_2$ are low-rank matrices and if the subspaces $E\mathcal{B}_k(A, \bar{C}_1)$ and $E\mathcal{B}_k(B, \bar{C}_2)$ have good approximation properties. Our approach consists of applying Algorithm 1 directly to the generalized Sylvester equation (3). Therefore, we now derive certain approximation properties of the solution to (3) that naturally suggest a proper choice of starting blocks. The low rank of the starting blocks will rely on the low-rank commutation property of the coefficients (4). Our reasoning can be described as follows:

- The solution to the generalized Sylvester Equation (3) can be represented as a converging Neumann series (5). By truncating this series, $X^{(\ell)}$ gives an approximation to the solution to (3), in the sense of Remark 1.
- The coefficients of the Neumann series (6) satisfy a sequence of Sylvester equations, where for each Sylvester equation, the right-hand side depends on the solution to the previous Sylvester equation. We consider approximate solutions to this sequence. More precisely, let $\tilde{Y}_j$ be the result of Algorithm 1 (as in Remark 3) applied to each Sylvester equation (6).
- The matrix $\tilde{X}^{(\ell)} = \sum_{j=0}^{\ell} \tilde{Y}_j$, which can be viewed as an approximation to the solution of (3), can be factorized as $\tilde{X}^{(\ell)} = LR^T$ such that $\text{Range}(L) \subseteq E\mathcal{B}_k(A, \bar{C}_1)$ and $\text{Range}(R) \subseteq E\mathcal{B}_k(B, \bar{C}_2)$. We give a characterization and a procedure for computing $\bar{C}_1$ and $\bar{C}_2$. One condition for these matrices to be low rank concerns the commutators (4) being low rank. These two matrices will be used as starting blocks in Algorithm 1.

Although the above reasoning is based on solving a sequence of Sylvester equations, our approach consists of applying Algorithm 1, only one time, directly to the generalized Sylvester equation (3).

We first need a technical result that shows that if the commutator of two matrices has low rank, then the corresponding commutator, where one matrix is taken to a given power, has also low rank. The rank increases at most linearly with respect to the power of the matrix. The precise statement is presented in the following lemma.

Lemma 1. *Suppose that A and N are matrices such that $[A, N] = U\tilde{U}^T$. Then,*

$$[A^j, N] = \sum_{k=0}^{j-1} A^k U \tilde{U}^T A^{j-k-1}.$$

Proof. The proof is by induction. The basis of induction is trivially verified for $j = 1$. Assume that the claim is valid for j; then, the induction step follows by observing that

$$[A^{j+1}, N] = A^{j+1}N - NA^{j+1} = A^j U\tilde{U}^T + (A^j N - NA^j)A$$

and by applying the induction hypothesis on $A^jN - NA^j$. □

As pointed out in Remark 3, C_1 and C_2 are natural starting blocks for the Sylvester equation. If we apply this result to the sequence of Sylvester equations in Theorem 1, with $\mathcal{L}$ and Π defined as in Equations 1–2, we obtain subspaces with a particular structure. For example, the approximation $L_0R_0^T$ to Y_0 provided by Algorithm 1 is such that $\text{Range}(L_0) \subseteq E\mathcal{B}_k(A, C_1)$ and $\text{Range}(R_0) \subseteq E\mathcal{B}_k(B, C_2)$. Because Y_0 is contained in the right-hand side of the definition of Y_1, in order to compute an approximation of Y_1, we should consider the subspaces $N_i \cdot E\mathcal{B}_k(A, C_1)$ and $M_i \cdot E\mathcal{B}_k(B, C_2)$ for $i = 1, \ldots, m$. By using the low-rank commutation property (4), such subspaces can be characterized by the following result.

Theorem 3. *Assume that $A \in \mathbb{R}^{n\times n}$ is nonsingular, and let $N \in \mathbb{R}^{n\times n}$ such that $[A, N] = U\tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n\times s}$. Let $C \in \mathbb{R}^{n\times r}$; then,*

$$N \cdot E\mathcal{B}_k(A, C) \subseteq E\mathcal{B}_k(A, (NC, U)).$$

Proof. Let $Np(A)Cw + Nq(A^{-1})Cv$ be a generator of $N \cdot E\mathcal{B}_k(A, C)$, where $p(x) = \sum_{j=0}^{k} \alpha_j x^j$ and $q(x) = \sum_{j=0}^{k} \beta_j x^j$. Then, with a direct usage of Lemma 1, the vector $Np(A)Cw$ can be expressed as an element of $E\mathcal{B}_k(A, (NC, U))$ in the following way:

$$Np(A)Cw = N\sum_{j=0}^{k}\alpha_j A^j Cw = p(A)NCw - \sum_{j=0}^{k}\sum_{\ell=0}^{j-1}\alpha_j A^\ell U(\tilde{U}^T A^{j-1-\ell} Cw).$$

We can show that $Nq(A^{-1})Cv$ belongs to the subspace $E\mathcal{B}_k(A, (NC, U))$ with the same procedure and by using that $[A^{-1}, N] = -(A^{-1}U)(A^{-T}\tilde{U})^T$. □

In order to ease the notation and improve the conciseness of the results that follow, we introduce the following multivariate generalization of the Krylov subspace for more matrices:

$$\mathcal{G}_d(N_1, \ldots, N_m; U) := \text{span}\left(\left\{p(N_1, \ldots, N_m)Uz | \deg(p) \le d, z \in \mathbb{R}^r\right\}\right),$$

where $U \in \mathbb{R}^{n\times r}$ and p is a noncommutative multivariate polynomial in the free algebra $\mathbb{R} < x_1, \ldots, x_N >$ (in the sense of chapter 10 in the work of Berstel et al.[36]).

Observation 2. Observe that $\mathcal{G}_d(N_1, \ldots, N_m; U)$ is the space generated by the columns of the matrices obtained by multiplying (in any order) $s \le d$ matrices N_i and the matrix U. In particular, this space can be equivalently characterized as

$$\mathcal{G}_d(N_1, \ldots, N_m; U) = \text{span}\left(\left\{N_{i_1}\cdots N_{i_s}Uz | 1 \le i_j \le m, 0 \le s \le d, z \in \mathbb{R}^r\right\}\right).$$

This definition generalizes the definition of the standard Krylov subspace in the sense that $\mathcal{G}_d(N; U) = \mathcal{B}_d(N, U)$.

The solution strategy for (3) outlined at the beginning of this subsection is formalized in the following theorem. In order to state the theorem, we need the result of the application of the extended Krylov method to the (standard) Sylvester equations of the form

$$A\mathcal{Y} + \mathcal{Y}B^T = C_1C_2^T, \tag{22a}$$

$$A\mathcal{Y} + \mathcal{Y}B^T = -\sum_{i=1}^{m}(N_iL_j)(M_iR_j)^T, \tag{22b}$$

as described in Simoncini,[5] and Breiten et al.[7] As already stated in Remark 3, this is identical to applying Algorithm 1 with $m = 0$.

Theorem 4. *Consider the generalized Sylvester equation* (3)*, with coefficients commuting according to* (4)*. Let* $\tilde{Y}_0 = L_0R_0^T$ *be the result of Algorithm 1 applied to the (standard) Sylvester equation* (22a) *with starting blocks* $\bar{C}_1 = C_1$ *and* $\bar{C}_2 = C_2$*. Moreover, for* $j = 0, \ldots, \ell - 1$*, let* $\tilde{Y}_{j+1} = L_{j+1}R_{j+1}^T$ *be the result of Algorithm 1 applied to the Sylvester equation* (22b) *with starting blocks* $\bar{C}_1 = (N_1L_j, \ldots, N_mL_j)$ *and* $\bar{C}_2 = (M_1R_j, \ldots, M_mR_j)$*. Let* $\tilde{X}^{(\ell)}$ *be the approximation of the truncated Neumann series* (7) *given by*

$$\tilde{X}^{(\ell)} := \sum_{j=0}^{\ell}\tilde{Y}_j.$$

Then, there exist matrices $L, R, \hat{C}_1^{(\ell)}, \hat{C}_2^{(\ell)}$ *such that* $\text{Range}(L) \subseteq E\mathcal{B}_{(\ell+1)d}(A, \hat{C}_1^{(\ell)})$ *and* $\text{Range}(R) \subseteq E\mathcal{B}_{(\ell+1)d}(B, \hat{C}_2^{(\ell)})$ *and*

$$\tilde{X}^{(\ell)} = LR^T,$$

where

$$\text{Range}\left(\hat{C}_1^{(\ell)}\right) \subseteq \quad \mathcal{G}_\ell(N_1, \ldots, N_m; C_1) + \mathcal{G}_{\ell-1}(N_1, \ldots, N_m; U), \tag{23a}$$

$$\text{Range}\left(\hat{C}_2^{(\ell)}\right) \subseteq \quad \mathcal{G}_\ell(M_1, \ldots, M_m; C_2) + \mathcal{G}_{\ell-1}(M_1, \ldots, M_m; Q), \tag{23b}$$

and $U := (U_1, \ldots, U_m)$*,* $Q := (Q_1, \ldots, Q_m)$*.*

Proof. We start proving that for $j = 0, \ldots, \ell$, there exists a matrix S_j such that $\text{Range}(L_j) \subseteq E\mathcal{B}_{(j+1)d}(A, S_j)$ and

$$\text{Range}(S_j) \subseteq \text{span}\left(\left\{\left(\prod_{k=1}^{j}N_{i_k}\right)C_1w + p(N_1, \ldots, N_m)\,Uz \,\middle|\, w \in \mathbb{R}^r, z \in \mathbb{R}^s, 1 \le i_k \le m, \deg(p) \le j-1\right\}\right), \tag{24}$$

where $s = \sum_{i=1}^{m} s_i$, and s_i denotes the number of columns of U_i. We prove this claim by induction. The basis of induction is trivially verified with $S_0 := C_1$ and using Remark 3. We now assume that the claim is valid for j and we

perform the induction step. Remark 3 implies that $\text{Range}(L_{j+1}) \subseteq E\mathcal{B}_d(A, (N_1 L_j, \ldots, N_m L_j))$. From Theorem 3 and the induction hypothesis, we have that $\text{Range}(N_i L_j) \subseteq E\mathcal{B}_{(j+1)d}(A, (N_i, S_j U_i))$ for any $i = 1, \ldots, m$. Therefore, we have that $\text{Range}(L_{j+1}) \subseteq E\mathcal{B}_{(j+2)d}(A, (N_1 S_j, \ldots, N_m S_j, U))$. We define $S_{j+1} := (N_1 S_j, \ldots, N_m S_j, U)$, which concludes the induction.

From (24), we now obtain the relation

$$\text{Range}((S_0, \ldots, S_j)) \subseteq \mathcal{G}_j(N_1, \ldots, N_m; C_1) + \mathcal{G}_{j-1}(N_1, \ldots, N_m; U)$$

that directly implies (23a) by setting $\hat{C}_1^{(\ell)} := (S_0, \ldots, S_\ell)$. Equation (23b) follows from completely analogous reasoning. The final conclusion follows by defining $L := (L_0, \ldots, L_\ell)$ and $R := (R_0, \ldots, R_\ell)$. □

The main message of the previous theorem can be summarized as follows. The low-rank factors of the approximation of $X^{(\ell)}$ (7) obtained by solving the Sylvester equations (6) with K-PIK[5,7] (that is equivalent to Algorithm 1 as discussed in Remark 3) are contained in an extended Krylov subspace with a specific choice of the starting blocks. In particular, the starting blocks are selected as $\bar{C}_1 = \hat{C}_1^{(\ell)}$, $\bar{C}_2 = \hat{C}_2^{(\ell)}$, where $\hat{C}_1^{(\ell)}$ and $\hat{C}_2^{(\ell)}$ fulfill (23a)–(23b). Our approach consists of applying Algorithm 1 directly to the generalized Sylvester equation (3) with this choice of the starting blocks.

A practical procedure that generates starting blocks that fulfill (23) consists of selecting $\bar{C}_1$ and $\bar{C}_2$ such that their columns are respectively a basis of the subspaces $\mathcal{G}_\ell(N_1, \ldots, N_m; C_1) + \mathcal{G}_{\ell-1}(N_1, \ldots, N_m; U)$ and $\mathcal{G}_\ell(M_1, \ldots, M_m; C_2) + \mathcal{G}_{\ell-1}(M_1, \ldots, M_m; Q)$. A basis of such spaces can be computed by using Observation 2. For example, a basis of $\mathcal{G}_2(N_1, N_2; U)$ can be obtained from the columns of the matrix

$$\left(U,\ N_1 U,\ N_2 U,\ N_1 N_2 U,\ N_2 N_1 U,\ N_1^2 U,\ N_2^2 U\right).$$

Observation 3. The choice of the starting blocks involves the parameter ℓ. In theory, a suitable choice of ℓ could be derived by using Remark 1. However, this is not always possible because the quantity $\rho(\mathcal{L}^{-1}\Pi)$ is, in many cases, not known and computationally demanding to compute/estimate. The choice of ℓ is a trade-off between accuracy and efficiency. The starting blocks $\hat{C}_1^{(\ell)}$ and $\hat{C}_2^{(\ell)}$, for large ℓ, provide spaces $E\mathcal{B}_k(A, \bar{C}_1)$ and $E\mathcal{B}_k(B, \bar{C}_2)$ with better approximation features but with potentially higher dimensions. See Figure 1. This leads to an increment in the computational cost of the whole procedure and to a more accurate approximation to the solution to (3).

Our approach is computationally attractive only if the starting blocks $\bar{C}_1 = \hat{C}_1^{(\ell)}$ and $\bar{C}_2 = \hat{C}_2^{(\ell)}$ have low rank. This condition is fulfilled if the commutators (4) are low-rank matrices; see Observation 2. Under this assumption, the advantages of the proposed method can be summarized as follows: Algorithm 1 generates only one pair of extended Krylov subspaces with given starting blocks. There are other methods based on generating several projection subspaces (with the same coefficient matrix), for example, a direct computation of $\tilde{X}^{(\ell)}$ or Damm,[8] Shank et al.[22] and section 5.3 of Benner et al.[37] An advantage of our approach with respect to these methods consists of avoiding redundancy in the approximation spaces. In particular, if several Krylov subspaces with the same coefficient matrix are generated independently of each other, they may have a nontrivial intersection or in general may have similar approximation properties. From a computational point of view, this means that considerable efforts are wasted to breed similar information.

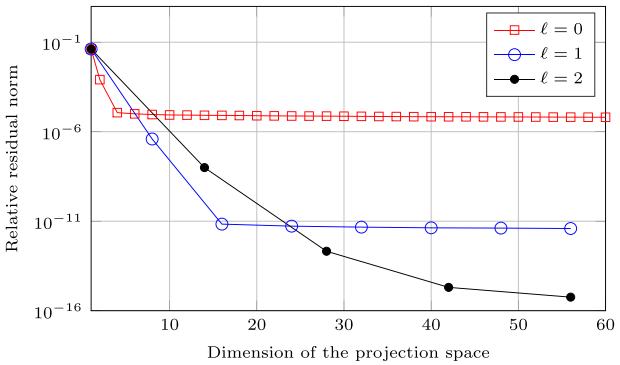

FIGURE 1 Convergence history of Algorithm 1 applied to a (randomly generated) generalized Lyapunov equation $AX + XA + NXN = cc^T$ with A circulant and N sum of a circulant matrix plus a rank one uu^T correction. The algorithm is tested for the stating blocks $\hat{C}^{(\ell)}$ with $\ell = 0, 1, 2$ selected according to Theorem 3, that is, $\hat{C}^{(0)} = c$, $\hat{C}^{(1)} = (c, Nc, u)$, $\hat{C}^{(2)} = (c, Nc, N^2 c, u, Nu)$

In certain cases, the dimension of the subspaces $\mathcal{G}_\ell$ is bounded for all the ℓ, that is, there exist matrices $\bar{C}_1 \in \mathbb{R}^{n\times \bar{r}_1}$ and $\bar{C}_2 \in \mathbb{R}^{n\times \bar{r}_2}$ such that Range($\hat{C}_1^{(\ell)}$) $\subseteq$ Range($\bar{C}_1$) and Range($\hat{C}_2^{(\ell)}$) $\subseteq$ Range($\bar{C}_2$) for all ℓ. This condition is satisfied, for example, if the matrix coefficients N_i, M_i are nilpotent/idempotent or, in general, if they have low degree minimal polynomials. Therefore, it is possible to select the starting blocks such that Algorithm 1 provides an approximation of $X^{(\ell)}$ for all ℓ, that is, the full series (5) is approximated. These situations naturally appear in applications; see the numerical example in Section 4.3.

3.3 | Krylov subspace method and low-rank matrices

Our numerical method can be improved for the following special case. We now consider a generalized Sylvester equation (3) where $N_i = \mathcal{U}_i\tilde{\mathcal{U}}_i^T$ and $M_i = \mathcal{Q}_i\tilde{\mathcal{Q}}_i^T$ are low-rank matrices. Obviously, the commutators $[A, N_i]$ and $[B, M_i]$ also have low rank, and the theory and the procedure presented in the previous section cover this case. However, the solution to (3) can be further characterized, and an efficient (and different) choice of the starting blocks $\bar{C}_1, \bar{C}_2$ can be derived. The assumption $\rho(\mathcal{L}^{-1}\Pi) < 1$ is no longer needed in order to justify the low-rank approximability. This property can be illustrated with a Sherman–Morrison–Woodbury argument as proposed by Benner et al.[4] The following proposition shows that the generalized Sylvester equation (3) can be implicitly written as a Sylvester equation with a right-hand side involving the columns of the matrices $\mathcal{U}_i$ and $\mathcal{Q}_i$ for $i = 1, \dots, m$. By using Remark 3, this leads to the following natural choice of starting blocks: $\bar{C}_1 = (C_1, \mathcal{U}_1, \dots, \mathcal{U}_m)$ and $\bar{C}_2 = (C_2, \mathcal{Q}_1, \dots, \mathcal{Q}_m)$.

Proposition 2. *Consider the generalized Sylvester equation* (3). *Assume that $\mathcal{L}$ is invertible and that $N_i = \mathcal{U}_i\tilde{\mathcal{U}}_i^T$ and $M_i = \mathcal{Q}_i\tilde{\mathcal{Q}}_i^T$ are such that $\mathcal{U}_i, \tilde{\mathcal{U}}_i \in \mathbb{R}^{n\times s_i}$ and $\mathcal{Q}_i, \tilde{\mathcal{Q}}_i \in \mathbb{R}^{n\times t_i}$. Then, there exist $\alpha_{i,j,\ell} \in \mathbb{R}$ for $j = 1, \dots, s_i$, $\ell = 1, \dots, t_i$, and $i = 1, \dots, m$, such that*

$$AX + XB^T = C_1C_2^T - \sum_{i,j,\ell} \alpha_{i,j,\ell}\, \mathcal{U}_i^{(j)}\mathcal{Q}_i^{(\ell)^T},$$

where $\mathcal{U}_i^{(j)}$ is the jth column of $\mathcal{U}_i$, and $\mathcal{Q}_i^{(\ell)}$ is the ℓth column of $\mathcal{Q}_i$.

Proof. The proof follows by theorem 4.1 of Ringh et al.,[9] proposition 3.3 of Damm,[8] or theorem 2.2 of Richter et al.[28] □

3.4 | Solving the projected problem

In order to apply Algorithm 1, we need to solve the projected problem in Step 4. The projected problem has to be solved in every iteration, and efficiency is therefore required in practice. For completeness, we now derive a procedure to solve the projected problem based on the Neumann series expansion derived in Section 2.1, although this is certainly not the only option. The derivation is based on the following observations. The projected problem is a small generalized Sylvester equation (3), and the computation of $X^{(\ell)}$ in (7) requires solving $\ell + 1$ Sylvester equations (6). Because the Sylvester equations (6) are defined by the same coefficients, they can be simultaneously reduced to triangular form (cf. section 3 of Richter et al.[28]) as follows:

$$U_A\widetilde{Y}_0 + \widetilde{Y}_0U_B^T = \widetilde{C}_1\widetilde{C}_2^T, \tag{25a}$$

$$U_A\widetilde{Y}_{j+1} + \widetilde{Y}_{j+1}U_B^T = -\sum_{i=1}^{m}\widetilde{N}_i\widetilde{Y}_j\widetilde{M}_i^T, \qquad j = 0, \dots, \ell - 1, \tag{25b}$$

where we have defined

$$\widetilde{C}_1 := Q_A^TC_1, \quad \widetilde{C}_2 := Q_B^TC_2, \quad \widetilde{N}_i := Q_A^TN_iQ_A, \quad \widetilde{M}_i := Q_B^TM_iQ_B, \tag{26}$$

and $A = Q_AU_AQ_A^T$ and $B = Q_BU_BQ_B^T$ denote the Schur decompositions. The Sylvester equations (25) with triangular coefficients can be efficiently solved with backward substitution as in the Bartels–Stewart algorithm,[38] and it holds that $X^{(\ell)} = Q_A(\sum_{j=0}^{\ell}\widetilde{Y}_j)Q_B^T$. The Frobenius norm of the residual $\mathcal{R}^{(\ell)} := AX^{(\ell)} + X^{(\ell)}B^T + \sum_{i=1}^{m}N_iX^{(\ell)}M_i^T - C_1C_2^T$ can be computed without explicitly constructing $X^{(\ell)}$ as follows:

$$\|\mathcal{R}^{(\ell)}\|_F = \left\|\sum_{i=1}^{m}\widetilde{N}_i\widetilde{Y}_\ell\widetilde{M}_i^T\right\|_F. \tag{27}$$

The previous relation follows by simply using the properties of the Frobenius norm (invariance under orthogonal transformations) and the relations (25).

In conclusion, the following iterative procedure can be used to approximate the solution to (3). The matrices (26) are precomputed; then, the Sylvester equations in triangular form (25) are solved until the residual of the Neumann series (27) is sufficiently small. The approximation $X^{(\ell)}$ is not computed during the iteration and only constructed after the iteration has completed. The procedure is summarized in Algorithm 2.

Algorithm 2: Neumann series approach for (3).

input : Matrix coefficients: $A, B, N_1 \dots, N_m, M_1, \dots, M_m, C_1, C_2$
output: Truncated Neumann series $X^{(\ell)}$

1 Compute the Schur decompositions $A = Q_A U_A Q_A^T$, $B = Q_B U_B Q_B^T$
2 Compute $\widetilde{C}_1$, $\widetilde{C}_2$, $\widetilde{N}_i$ $\widetilde{M}_i$ for all $i = 1, \dots, m$ according to (26)
3 Solve $U_A \widetilde{Y}_0 + \widetilde{Y}_0 U_B^T = \widetilde{C}_1 \widetilde{C}_2^T$ and set $\widetilde{X} = \widetilde{Y}_0$
for $j = 0, 1, \dots$ *till convergence*, **do**
4 Solve $U_A \widetilde{Y}_{j+1} + \widetilde{Y}_{j+1} U_B^T = -\sum_{i=1}^m \widetilde{N}_i \widetilde{Y}_j \widetilde{M}_i^T$ and set $\widetilde{X} = \widetilde{X} + \widetilde{Y}_{j+1}$
5 Compute $\|\mathcal{R}^{(j+1)}\|_F = \| \sum_{i=1}^m \widetilde{N}_i \widetilde{Y}_{j+1} \widetilde{M}_i^T \|_F$
if $\|\mathcal{R}^{(j+1)}\|_F \leq$ `tol`, **then**
6 Set $\ell = j + 1$
7 Break
end
end
8 Return $X^{(\ell)} = Q_A \widetilde{X} Q_B^T$

4 | NUMERICAL EXAMPLES

We now illustrate our approach with several examples. In the first two examples, we compare our approach with two different methods for generalized Lyapunov equations: BilADI[4] and Generalized Lyapunov Extended Krylov (GLEK) method.[22] The results are generally in favor of our approach, because the other methods are less specialized to the specific structure. However, they have a wider applicable problem domain. Two variants of BilADI are considered. In the first variant, we select the Wachspress shifts[39] computed with the software available on Saak's web page.* In the second variant, $\mathcal{H}_2$-optimal shifts[37] are used. The GLEK code is available at the web page of Simoncini.† This algorithm requires fine-tuning several thresholds. We selected `tol_inexact`$= 10^{-2}$, whereas the default setting is used for all the other thresholds. The implementation of our approach is based on the modification of K-PIK[5,7] for generalized Sylvester equation as described in Algorithm 1. The projected problems, computed in Step 4, are solved with the procedure described in the Section 3.4. A MATLAB implementation of Algorithm 1 is available online.‡

In all the methods that we test, the stopping criterion is based on the relative residual norm, and the algorithms are stopped when it reaches tol $= 10^{-6}$. We compare the number of iterations, memory requirements, rank of the computed approximation, number of linear solves (involving the matrices A and B potentially shifted), and total execution CPU times.

As memory requirement (denoted Mem. in the following tables), we consider the number of vectors of length n stored during the solution process. In particular, for Algorithm 1, it consists of the dimension of the approximation space. In GLEK, a sequence of extended Krylov subspaces is generated, and the memory requirement corresponds to the dimension of the largest space in the sequence. For the bilinear ADI approach, the memory requirement consists of the number of columns of the low-rank factor of the solution. For GLEK, we just report the number of outer iterations. The CPU times reported for BilADI do not take into account the time for the shifts computation. For the linear solves, the LU factors are precomputed and reused in the algorithms. All results were obtained with MATLAB R2015a on a computer with two 2 GHz processors and 128 GB of RAM.

*https://www2.mpi-magdeburg.mpg.de/mpcsc/mitarbeiter/saak/Software/adipars.php
†http://www.dm.unibo.it/~simoncin/software.html
‡http://www.dm.unibo.it/~davide.palitta3

TABLE 1 Multiple input multiple output system example

	γ	Its.	Mem.	rank(X)	Lin. solves	CPU time
BilADI (4 Wach.)	1/6	10	55	55	320	51.26
BilADI (8 $\mathcal{H}_2$-opt.)	1/6	10	55	55	320	51.54
GLEK	1/6	9	151	34	644	14.17
Algorithm 1	1/6	6	72	60	36	3.77
BilADI (4 Wach.)	1/5	14	71	71	588	55.15
BilADI (8 $\mathcal{H}_2$-opt.)	1/5	14	69	69	586	54.31
GLEK	1/5	12	173	39	1,016	22.06
Algorithm 1	1/5	6	72	61	36	4.23
BilADI (4 Wach.)	1/4	24	89	89	1,454	67.61
BilADI (8 $\mathcal{H}_2$-opt.)	1/4	23	89	89	1,371	66.83
GLEK	1/4	21	218	50	2,348	51.49
Algorithm 1	1/4	8	96	81	48	6.72

Note. Comparison of low-rank methods for $n = 50{,}000$. Mem. = memory requirement.

4.1 | A multiple input multiple output system

The time invariant multiple input multiple output bilinear system described in example 2 of Lin et al.[40] yields the following generalized Lyapunov equation:

$$AX + XA^T + \gamma^2 \sum_{i=1}^{2} N_i X N_i^T = CC^T, \tag{28}$$

where $\gamma \in \mathbb{R}$, $\gamma > 0$, $A = \text{tridiag}(2, -5, 2)$, $N_1 = \text{tridiag}(3, 0, -3)$, and $N_2 = -N_1 + I$. We consider $C \in \mathbb{R}^{n\times 2}$ being a normalized random matrix. In the context of bilinear systems, the solution to (28), referred to as *Gramian*, is used for computing energy estimates and reachability of the states. The number γ is a scaling parameter selected in order to ensure the solvability of the problem (28) and the positive definiteness of the solution, namely, $\rho(\mathcal{L}^{-1}\Pi) < 1$. This parameter corresponds to rescaling the input of the underlying problem with a possible reduction in the region where energy estimates hold. Therefore, it is preferable not to employ very small values of γ. See Benner et al.[2] for detailed discussions.

For this problem, the commutators have low rank, more precisely $[A, N_1] = -[A, N_2] = U\tilde{U}^T$, with $U = 2\sqrt{3}(e_1, e_n)$ and $\tilde{U} = 2\sqrt{3}(e_1, -e_n)$. As proposed in Section 3.2, we use Algorithm 1 with starting blocks $\bar{C}_1 = \bar{C}_2 = (C, N_1C, U)$ because $\text{Range}(C_1^{(1)}) = \text{Range}((C, N_1C, N_2C, U)) = \text{Range}\,(C, N_1C, U)$. Table 1 illustrates the performances of our approach and other low-rank methods, GLEK and the BilADI, as γ varies. We notice that the number of linear solves that our projection method requires is always much less than for the other methods. Moreover, it seems that moderate variations of γ, which correspond to variations of $\rho(\mathcal{L}^{-1}\Pi)$, have a smaller influence on the number of iterations in our method compared with other algorithms.

4.2 | A low-rank problem

We now consider the following generalized Lyapunov equation:

$$AX + XA^T + UV^TXVU^T = cc^T, \tag{29}$$

where $A = n^2\text{tridiag}(1, -2, 1)$ and $U, V \in \mathbb{R}^{n\times m}$, $c \in \mathbb{R}^n$ have random entries and unit norm. We use Algorithm 1, and as proposed in Section 3.3, we select $\bar{C}_1 = \bar{C}_2 = (c, U)$ as starting blocks. In Table 2, we report the results of the comparison with other methods for $m = 1$. We notice that our approach requires the lowest number of linear solves. The ADI approaches demand the lowest storage because of the column compression strategy performed at each iteration. However, due to the large number of linear solves, these methods are slower compared with our approach. For large-scale problems, the BilADI method with 4 Wachspress shifts does not converge in 500 iterations. GLEK provides the solution with the smallest rank.

We now consider (29) for $m > 1$. Notice that, in Equation (3), this corresponds to have the operator Π (2) defined by the sum of m terms of rank 1. In particular, we apply Algorithm 1 to Equation (29) for $m = 5, 10, 15$. The results are collected

TABLE 2 Comparison of low-rank methods applied to (29) varying n with $m = 1$

	n	Its.	Mem.	rank(X)	Lin. solves	CPU time
BilADI (4 Wach.)	10,000	60	57	57	2,462	4.25
BilADI (8 $\mathcal{H}_2$-opt.)	10,000	42	55	55	1,420	2.54
GLEK	10,000	4	240	28	310	3.10
Algorithm 1	10,000	46	184	49	92	2.77
BilADI (4 Wach.)	50,000	327	61	61	18,673	315.56
BilADI (8 $\mathcal{H}_2$-opt.)	50,000	96	61	61	4,580	81.47
GLEK	50,000	4	454	28	565	24.78
Algorithm 1	50,000	78	312	47	156	21.09
BilADI (4 Wach.)	100,000	–	–	–	–	–
BilADI (8 $\mathcal{H}_2$-opt.)	100,000	84	65	65	4,058	174.04
GLEK	100,000	4	457	29	631	66.77
Algorithm 1	100,000	97	388	44	194	55.58

Note. Mem. = memory requirement.

TABLE 3 Algorithm 1 applied to (29) varying n and m

n	m	Its.	Mem.	rank(X)	Lin. solves	CPU time
10,000	5	33	396	50	198	9.38
10,000	10	27	594	48	297	19.87
10,000	15	24	768	44	384	27.35
50,000	5	55	660	43	330	54.87
50,000	10	45	990	41	495	117.26
50,000	15	40	1,280	42	640	245.87
100,000	5	68	816	43	408	133.72
100,000	10	56	1,232	41	616	332.68
100,000	15	50	1,600	44	800	743.86

Note. Mem. = memory requirement.

in Table 3. The number of iterations needed decreases as m increases. However, because the rank of the starting block increases with m, the dimension of the approximation space increases and thus the number of linear solves. As a result, the computation time increases with m.

If we replace the matrix A with A/n^2 in Equation (29), neither BilADI nor GLEK converges because the Lyapunov operator is no longer dominant, that is, $\rho(\mathcal{L}^{-1}\Pi) > 1$. However, our algorithm still converges, and for $n = 10,000$, $m = 1$, it provides a solution X in 46 iterations with rank(X)=184. In this case, the projected problems cannot be solved with the approach described in the Section 3.4. However, because the projected problems are also of the form (29), they can be solved with a Sherman–Morrison–Woodbury approach for the matrix equation.[8,9,28] In this case, we used the method presented in section 3 of Damm.[8]

4.3 | Inhomogeneous Helmholtz equation

In the last example, we analyze the complexity of Algorithm 1 when solving a large-scale generalized Sylvester equation stemming from a finite difference discretization of a PDE. More precisely, we consider the following inhomogeneous Helmholtz equation:

$$\begin{cases} -\Delta u(x,y) + \kappa(x,y)u(x,y) = f(x,y), \quad (x,y) \in [0,1] \times \mathbb{R}, \\ u(0,y) = u(1,y) = 0, \\ u(x,y+1) = u(x,y). \end{cases} \tag{30}$$

The boundary conditions are periodic in the y-direction and homogeneous-Dirichlet in the x-direction. The wavenumber $\kappa(x,y)$ and the forcing term $f(x,y)$ are 1-periodic functions in the y-direction. In particular, they are respectively the periodic extensions of the scaled indicator functions $\chi_{[1/2,1]^2}$ and $100\chi_{[1/4,1/2]^2}$. The discretization of Equation (30) with the finite difference method, using n nodes multiple of four, leads to the following generalized Sylvester equation:

$$AX + XB^T + NXN^T = CC^T, \tag{31}$$

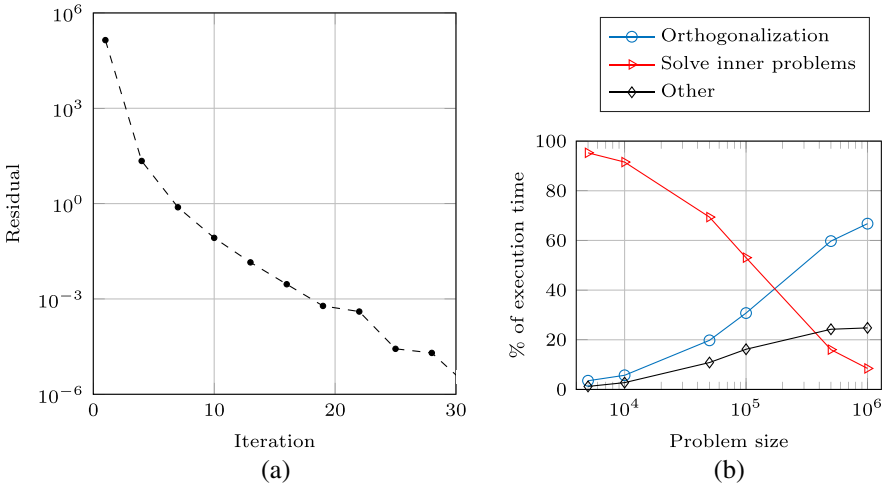

FIGURE 2 Simulations for the inhomogeneous Helmholtz equation. (a) Residual norm history for problem size $n = 10,000$. (b) CPU time (%) of the main parts of Algorithm 1 with $d = 30$. Total time (in seconds) are 23.71, 25.41, 34.33, 48.19, 163.34, and 315.74, respectively

where $B = -\text{tridiag}(1, -2, 1)/h^2$, $h = 1/(n-1)$ is the mesh size, $A = B - (e_1, e_n)(e_n, e_1)^T/h^2$, and

$$N = \begin{pmatrix} O_{n/2} & O_{n/2} \\ O_{n/2} & I_{n/2} \end{pmatrix} \in \mathbb{R}^{n\times n}, \quad C = (c_1, \ldots, c_n)^T, \; c_i = \begin{cases} 10, & \text{if } i \in [n/4, n/2], \\ 0, & \text{otherwise.} \end{cases}$$

A direct computation shows that $[A, N] = U\tilde{U}^T$ and $[B, N] = Q\tilde{Q}^T$, where

$$U = n(e_{n/2+1}, e_{n/2}, e_1, e_n), \qquad \tilde{U} = n(e_{n/2}, -e_{n/2+1}, -e_n, e_1),$$
$$Q = n(e_{n/2+1}, e_{n/2}), \qquad \tilde{Q} = n(e_{n/2}, -e_{n/2+1}).$$

Algorithm 1 is not applicable to Equation (31) because the matrix A is singular. However, in our approach, it is possible to shift the Sylvester operator. In particular, we can rewrite Equation (31) as

$$(A + I)X + XB^T + NXN^T - X = CC^T.$$

It is now possible to apply Algorithm 1 because $A + I$ is nonsingular. For this problem, it holds $N^2 = N$ and then $\mathcal{G}_\ell(N, I; C) = \text{Range}((C, NC))$ for all $\ell \geq 1$. We now note that $[A + I, N] = [A, N]$ and that $NC = 0$ and $\text{Range}((U, NU)) = \text{Range}(U)$. Hence, according to Theorem 3, we select $\tilde{C}_1 = (C, U)$ and $\tilde{C}_2 = (C, Q)$ as starting blocks. Notice that, with this choice, Algorithm 1 provides an approximation of $X^{(\ell)}$ for every $\ell \geq 0$. We fix the number of iterations $d = 30$ in Algorithm 1, and we vary the problem size n. In Figure 2B, we report the percentages of the overall execution time devoted to the orthogonalization procedure (Steps 9–11), to the solution of the inner problems (Step 4) and to the remaining steps of the algorithm. We can see that for very large problems, most of the computational effort is dedicated to the orthogonalization procedure. See Figure 2A for an illustration of the converge history for the problem of size $n = 10,000$.

5 | CONCLUSIONS AND OUTLOOK

The method that we have proposed for solving (3) is directly based on the low-rank commutation feature of the matrix coefficients (4). We have applied and adapted our procedure to problems in control theory and discretization of PDEs that naturally present this property. The structured matrices that present this feature are already analyzed in literature, although, to our knowledge, this was never exploited in the setting of Krylov-like methods for matrix equations. Low-rank commuting matrices are usually studied with the *displacement operators*. More precisely, for a given matrix Z, the displacement operator is defined as $F(A) := AZ - ZA$. For many specific choices of the matrix Z (e.g., Jordan block and circulant), it is possible to characterize the displacement operator and describe the matrices that are low rank commuting with Z; see, for example, Kailath and Sayed,[41] Beckermann and Townsend,[42] and chapter 2, section 11 of Bini et al.[43] and

the references therein. The theory concerning the displacement operator may potentially be used to classify the problems that can be solved with our approach.

The approach we have pursued in this paper is based on the extended Krylov subspace method. However, it seems possible to extend this to the rational Krylov subspace method[6] because the commutator $[A, N]$ is invariant under translations of the matrix A. Further research is needed to characterize the spaces and study efficient shift-selection strategies.

In each iteration of Algorithm 1, the residual can be computed without explicitly constructing the current approximation of the solution and by only using the solution of the projected problem. It may be possible to compute the residual norm even without explicitly solving the projected problems as proposed by Palitta et al.[44] for Lyapunov and Sylvester equations with symmetric matrix coefficients.

In conclusion, we wish to point out that the low-rank approximability characterization may be of use outside the scope of projection methods. For instance, the Riemannian optimization methods are designed to compute the best rank k approximation (in the sense of, e.g., Kressner et al.,[20] and Vandereycken and Vandewalle[21]) to the solution of the matrix equation. This approach is effective only if k is small, that is, the solution is approximable by a low-rank matrix, for which we have provided sufficient conditions.

ACKNOWLEDGEMENT

We wish to thank Tobias Breiten (University of Graz) for kindly providing the code that helped us implement BilADI[4] used in Section 4. We also thank Stephen D. Shank (Temple University) for providing us the GLEK code before its online publication.

This research commenced during a visit of the third author to the KTH Royal Institute of Technology. The warm hospitality received is greatly appreciated. The work of the third author is partially supported by INdAM-GNCS under the 2017 Project "Metodi numerici avanzati per equazioni e funzioni di matrici con struttura". The other authors gratefully acknowledge the support of the Swedish Research Council under Grant 621-2013-4640.

ORCID

Elias Jarlebring https://orcid.org/0000-0001-9443-8772
Giampaolo Mele http://orcid.org/0000-0002-6990-445X
Davide Palitta https://orcid.org/0000-0002-6987-4430
Emil Ringh http://orcid.org/0000-0001-6279-6145

REFERENCES

1. Simoncini V. Computational methods for linear matrix equations. SIAM Rev. 2016;58(3):377–441.
2. Benner P, Damm T. Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems. SIAM J Control Optim. 2011;49(2):686–711.
3. Lancaster P. Explicit solutions of linear matrix equations. SIAM Rev. 1970;12(4):544–566.
4. Benner P, Breiten T. Low rank methods for a class of generalized Lyapunov equations and related issues. Numer Math. 2013;124(3):441–470.
5. Simoncini V. A new iterative method for solving large-scale Lyapunov matrix equations. SIAM J Sci Comput. 2007;29(3):1268–1288.
6. Druskin V, Simoncini V. Adaptive rational Krylov subspaces for large-scale dynamical systems. Syst Control Lett. 2011;60(8):546–560.
7. Breiten T, Simoncini V, Stoll M. Low-rank solvers for fractional differential equations. Electron Trans Numer Anal. 2016;45:107–132.
8. Damm T. Direct methods and ADI-preconditioned Krylov subspace methods for generalized Lyapunov equations. Numer Linear Algebra Appl. 2008;15(9):853–871.
9. Ringh E, Mele G, Karlsson J, Jarlebring E. Sylvester-based preconditioning for the waveguide eigenvalue problem. Linear Algebra Appl. 2018;542:441–63.
10. Powell CE, Silvester D, Simoncini V. An efficient reduced basis solver for stochastic Galerkin matrix equations. SIAM J Sci Comput. 2017;39(1):A141–A163.
11. Palitta D, Simoncini V. Matrix-equation-based strategies for convection–diffusion equations. BIT Numer Math. 2016;56(2):751–776.
12. Grasedyck L. Existence of a low rank or $\mathcal{H}$-matrix approximant to the solution of a Sylvester equation. Numer Linear Algebra Appl. 2004;11(4):371–389.
13. Baker J, Embree M, Sabino J. Fast singular value decay for Lyapunov solutions with nonnormal coefficients. SIAM J Matrix Anal Appl. 2015;36(2):656–668.

14. Grasedyck L. Existence and computation of low Kronecker-rank approximations for large linear systems of tensor product structure. Computing. 2004;72(3):247–265.
15. Kressner D, Tobler C. Krylov subspace methods for linear systems with tensor product structure. SIAM J Matrix Anal Appl. 2010;31(4):1688–1714.
16. Benner P, Li RC, Truhar N. On the ADI method for Sylvester equations. J Comput Appl Math. 2009;233(4):1035–1045.
17. Benner P, Kürschner P. Computing real low-rank solutions of Sylvester equations by the factored ADI method. Comput Math Appl. 2014;67(9):1656–1672.
18. Baur U. Low rank solution of data-sparse Sylvester equations. Numer Linear Algebra Appl. 2008;15(9):837–851.
19. Baur U, Benner P. Factorized solution of Lyapunov equations based on hierarchical matrix arithmetic. Computing. 2006;78(3):211–234.
20. Kressner D, Steinlechner M, Vandereycken B. Preconditioned low-rank Riemannian optimization for linear systems with tensor product structure. SIAM J Sci Comput. 2016;38(4):A2018–A2044.
21. Vandereycken B, Vandewalle S. A Riemannian optimization approach for computing low-rank solutions of Lyapunov equations. SIAM J Matrix Anal Appl. 2010;31(5):2553–2579.
22. Shank SD, Simoncini V, Szyld DB. Efficient low-rank solution of generalized Lyapunov equations. Numer Math. 2016;134(2):327–342.
23. Kressner D, Sirković P. Truncated low-rank methods for solving general linear matrix equations. Numer Linear Algebra Appl. 2015;22(3):564–583.
24. Ballani J, Grasedyck L. A projection method to solve linear systems in tensor format. Numer Linear Algebra Appl. 2013;20(1):27–43.
25. Bouhamidi A, Jbilou K. A note on the numerical approximate solutions for generalized Sylvester matrix equations with applications. Appl Math Comput. 2008;206(2):687–694.
26. Li ZY, Zhou B, Wang Y, Duan GR. Numerical solution to linear matrix equation by finite steps iteration. IET Control Theory Appl. 1994;31(1):227–251.
27. Flagg G, Gugercin S, Multipoint Volterra series interpolation and H_2 optimal model reduction of bilinear systems. SIAM J Matrix Anal Appl. 2015;36(2):549–579.
28. Richter S, Davis LD, Collins JR EG. Efficient computation of the solutions to modified Lyapunov equations. SIAM J Matrix Anal Appl. 1993;14(2):420–431.
29. Zhang L, Lam J. On H_2 model reduction of bilinear systems. Automatica J IFAC. 2002;38(2):205–216.
30. Kato T. Perturbation theory for linear operators. Berlin, Germany: Springer-Verlag; 1995.
31. Jaimoukha IM, Kasenally EM. Krylov subspace methods for solving large Lyapunov equations. SIAM J Numer Anal. 1994;31(1):227–251.
32. Gutknecht MH. Block Krylov space methods for linear systems with multiple right-hand sides: An introduction. Modern mathematical models, methods and algorithms for real world systems. New Delhi, India: Anamaya Publishers; 2007.
33. Einstein E, Johnson CR, Lins B, Spitkovsky I. The ratio field of values. Linear Algebra Appl. 2011;434(4):1119–1136.
34. Simoncini V, Knizhnerman L. A new investigation of the extended Krylov subspace method for matrix function evaluations. Numer Linear Algebra Appl. 2010;17(4):615–638.
35. Frommer A, Lund K, Szyld DB. Block Krylov subspace methods for functions of matrices. Electron Trans Numer Anal. 2017;47:100–126.
36. Berstel J, Reutenauer C. Noncommutative rational series with applications, vol 137. Cambridge, UK: Cambridge University Press; 2011.
37. Benner P, Breiten T. Interpolation-based $\mathcal{H}_2$-model reduction of bilinear control systems. SIAM J Matrix Anal Appl. 2012;33(3):859–885.
38. Bartels RH, Stewart GW. Algorithm 432: solution of the matrix equation $AX + XB = C$. Comm ACM. 1972;15:820–826.
39. Wachspress E. The ADI model problem. New York, NY: Springer; 2013.
40. Lin Y, Bao L, Wei Y. Order reduction of bilinear MIMO dynamical systems using new block Krylov subspaces. Comput Math Appl. 2009;58(6):1093–1102.
41. Kailath T, Sayed AH. Displacement structure: theory and applications. SIAM Review. 1995;37(3):297–386.
42. Beckermann B, Townsend A. On the singular values of matrices with displacement structure. SIAM J Matrix Anal Appl. 2017;38(4):1227–1248.
43. Bini DA, Pan V. Polynomial and matrix computations: Fundamental algorithms. New York, NY: Springer Science Business Media; 2012.
44. Palitta D, Simoncini V. Computationally enhanced projection methods for symmetric Sylvester and Lyapunov equations. J Comput Appl Math. 2018;330:648–659.

How to cite this article: Jarlebring E, Mele G, Palitta D, Ringh E. Krylov methods for low-rank commuting generalized Sylvester equations. *Numer Linear Algebra Appl*. 2018;e2176. https://doi.org/10.1002/nla.2176