SIAM REVIEW
Vol. 58, No. 3, pp. 377–441

Computational Methods for Linear Matrix Equations*

V. Simoncini†

Abstract. Given the square matrices A, B, D, E and the matrix C of conforming dimensions, we consider the linear matrix equation $A\mathbf{X}E + D\mathbf{X}B = C$ in the unknown matrix $\mathbf{X}$. Our aim is to provide an overview of the major algorithmic developments that have taken place over the past few decades in the numerical solution of this and related problems, which are producing reliable numerical tools in the formulation and solution of advanced mathematical models in engineering and scientific computing.

Key words. Sylvester equation, Lyapunov equation, Stein equation, multiple right-hand side, generalized matrix equations, Schur decomposition, large scale computation

AMS subject classifications. 65F10, 65F30, 15A06

DOI. 10.1137/130912839

Contents

*Received by the editors March 13, 2013; accepted for publication (in revised form) August 19, 2015; published electronically August 4, 2016.
http://www.siam.org/journals/sirev/58-3/91283.html
†Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato 5, I-40127 Bologna, Italy, and IMATI-CNR, Pavia (valeria.simoncini@unibo.it).

1. Introduction.

Given the real or complex square matrices A, D, E, B and the matrix C of conforming dimensions, we consider the linear matrix equation

$$AXE + DXB = C \tag{1}$$

in the unknown matrix[1] $\mathbf{X}$, and its various generalizations. If E and D are identity matrices, then (1) is called the Sylvester equation, as its first appearance is usually associated with the work of J. J. Sylvester [240]; if in addition $B = A^*$, where A^* is the conjugate transpose of A, then the equation is called the Lyapunov equation in honor of A. M. Lyapunov and his early contributions to the stability problem of motion; see [14] and the entire issue of the same journal. We shall mainly consider the generic case, thus assuming that all the matrices involved are nonzero.

Under certain conditions on the coefficient matrices, (1) has a unique solution with available elegant and explicit closed forms. These are usually inappropriate as computational devices, either because they involve estimations of integrals, or because their computation can be polluted with numerical instabilities of various sorts. Nevertheless, closed forms and other properties of the solution matrix have strongly influenced the computational strategies that have led to most algorithms used today for numerically solving (1), in the case of small or large dimensions of the coefficient matrices. Due to the availability of robust and reliable core algorithms, (1) now arises in an increasingly larger number of scientific computations, from statistics to dynamical systems analysis, with a major role in control applications and also as a workhorse of more computationally intensive methods. In section 3 we will briefly review this broad range of numerical and application problems.

Our aim is to provide an overview of the major algorithmic developments that have taken place in the past few decades in the numerical solution of (1) and of related problems, both in the small and large scale cases. A distinctive feature in the large scale setting is that although the coefficient matrices may be sparse, the solution matrix is usually dense and thus impossible to store in memory. Therefore, ad hoc strategies need to be devised to approximate the exact solution in an affordable manner.

Functions related to the solution matrix $\mathbf{X}$, such as the spectrum, the trace, and the determinant, also have important roles in stability analysis and other applications. Although we shall not discuss in detail the computational aspects associated

[1] Here and in what follows we shall use boldface letters to denote the unknown solution matrices.

with these functions, we shall occasionally point to relevant results and appropriate references.

Linear matrix equations have received considerable attention since the early 1900s and were the topic of many elegant and thorough studies in the 1950s and 1960s, which used deep tools of matrix theory and functional analysis. The field continues to prosper with the analysis of new challenging extensions of the main equation (1), very often stimulated by application problems. Our contribution is intended to focus on the computational methods for solving these equations. For this reason, in our presentation we will mostly sacrifice the theoretical results, for which we refer the interested reader to, e.g., [90], [165], [131], [40].

The literature on the Lyapunov equation is particularly rich, due to the prominent role of this matrix equation in control theory. In particular, many authors have focused on numerical strategies specifically associated to this equation. As a consequence, the Sylvester and Lyapunov equations have somehow evolved differently. For these reasons, and to account for the literature in a homogeneous way, we shall first discuss numerical strategies for the Sylvester equation, and then treat in detail the Lyapunov problem. For A and B of size up to a few thousand, the Schur decomposition based algorithm by Bartels and Stewart [15] has since its appearance become the main numerical solution tool. In the large scale case, various directions have been taken and a selection of effective algorithms is available, from projection methods to sparse format iterations. Despite a lot of intense work in the past 15–20 years, the community has not entirely agreed upon the best approaches for all settings, hence the need for an overview that aims to analyze where the field stands at this point.

For A and B of the order of 10^4 or larger, the solution $\mathbf{X}$ cannot be stored explicitly; current memory-effective strategies rely on factored low-rank or sparse approximations. The possibility of computing a memory conserving good approximate solution in the large scale case depends highly on the data. In particular, for C full rank, accurate low-rank approximations may be hard, if not impossible, to find. For instance, the equation $A\mathbf{X} + \mathbf{X}A^* = I$ with A nonsingular and symmetric admits the unique solution $\mathbf{X} = \frac{1}{2}A^{-1}$, which is obviously full rank, with not necessarily quickly decreasing eigenvalues, so that a good low-rank approximation cannot be determined.

The distinction between small, moderate, and large size is clearly architecture-dependent. In what follows we shall refer to "small" and "medium" problem sizes when the coefficient matrices have dimensions of a few thousand at most; on high performance computers these dimensions can be considerably larger. Small and medium size linear equations can be solved with decomposition-based methods on laptops with moderate computational effort. The target for current large scale research is matrices of dimensions $\mathcal{O}(10^6)$ or larger, with a variety of sparsity patterns.

Throughout the article we shall assume either that E, D are the identity or that at least one of them is nonsingular. Singular E, D have great relevance in control applications associated with differential-algebraic equations and descriptor systems but require a specialized treatment, which can be found, for instance, in [164].

Equation (1) is a particular case of the linear matrix equation

$$A_1\mathbf{X}B_1 + A_2\mathbf{X}B_2 + \cdots + A_k\mathbf{X}B_k = C, \tag{2}$$

with A_i, B_i, $i = 1, \ldots, k$, square matrices and C of dimension $n \times m$. While up to 15–20 years ago this multiterm equation could be rightly considered to be of mainly theoretical interest, the recent developments associated with problems stemming from applications with parameters or a dominant stochastic component have brought multiterm linear matrix equations forward to play a fundamental role; see sections 3 and

7.2 for applications and references. Equation (2) is very difficult to analyze in its full generality, and necessary and sufficient conditions for the existence and uniqueness of the solution $\mathbf{X}$ explicitly based on $\{A_i\}, \{B_i\}$ are hard to find, except for some very special cases [165], [157]. While from a theoretical viewpoint the importance of taking into account the structure of the problem has been acknowledged [157], this has not been so for computational strategies, especially for large scale problems. The algorithmic device most commonly used for (2) consists in transforming the matrix equation above into a vector form by means of the Kronecker product (defined below). The problem of the efficient numerical solution of (2), with a target complexity of at most $\mathcal{O}(n^3 + m^3)$ operations, has only recently started to be addressed. The need for a low complexity method is particularly compelling whenever either or both A_i and B_i have large dimensions. Approaches based on the Kronecker formulations were abandoned for (1) as core methods, since algorithms with a complexity of a modest power of the coefficient matrices' dimension had become available. The efficient numerical solution to (2) thus represents the next frontier for linear matrix equations, to assist the rapidly developing application models.

Various generalizations of (1) have also been tackled in the literature, as they are more and more often encountered in applications. This is the case, for instance, for bilinear equations (in two unknown matrices) and for *systems* of bilinear equations. These are an open computational challenge, especially in the large scale case, and their efficient numerical solution would provide a great advantage for emerging mathematical models; we discuss these generalizations in section 7.

A very common situation arises when $B = 0$ and C is tall in (1), so that the matrix equation reduces to a standard linear system with multiple right-hand sides, the columns of C. This is an important problem in applications, and a significant body of literature is available, with a vast number of contributions made in the past thirty years, in particular in the large scale case, for which we refer the reader to [214] and to the more recent list of references in [113].

After a brief account in section 3 of the numerous application problems where linear matrix equations arise, we recall the main properties of these equations, together with possible explicit forms for their solution matrix. The rest of this article describes many approaches that have been proposed in the recent literature: we first treat the Sylvester equation when A and B are small, when one of the two is large, and when both are large. Indeed, rather different approaches can be employed depending on the size of the two matrices. We then focus on the Lyapunov equation: due to its relevance in control, many developments have specifically focused on this equation, therefore the problem deserves a separate treatment. We describe the algorithms that were specifically designed to take advantage of the symmetry, while only mentioning the solution methods that are common to the Sylvester equation. The small scale problem is computationally well understood, whereas the large scale case has seen quite significant developments made in the past ten years. Later sections report on the computational devices associated with the numerical solution of various generalizations of (1), which have been developed over the past few years.

2. Notation and Preliminary Definitions.

Unless stated otherwise, throughout the paper we shall assume that the coefficient matrices are real. Moreover, spec(A) denotes the set of eigenvalues of A, and $A^\top$, A^* denote the transpose and conjugate transpose of A, respectively. For $z \in \mathbb{C}$, $\bar{z}$ is the complex conjugate of z.

A matrix A is *stable* if all its eigenvalues have negative real part, and *negative definite* if for all unit 2-norm complex vectors x, the quantity $x^* A x$ has negative real part, namely, the field of values $W(A) = \{z \in \mathbb{C} : \ z = x^*Ax,\, x \in \mathbb{C}^n, \|x\| = 1\}$ is contained in the open left half complex plane. The notation $A \succ 0$ ($A \succeq 0$) states that A is a Hermitian and positive definite (semidefinite) matrix.

The vector e_i denotes the ith column of the identity matrix, whose dimension will be clear from the context; I_n denotes the identity matrix of size n, and the subscript will be omitted when clear from the context. Throughout, given $x \in \mathbb{C}^n$, $\|x\|$ denotes the 2-norm of x, $\|A\|$ or $\|A\|_2$ denotes the matrix norm induced by the vector 2-norm, while $\|A\|_F$ denotes the Frobenius norm of $A = (a_{i,j})_{i=1,\ldots,n,j=1,\ldots,m}$, that is, $\|A\|_F^2 = \sum_{i,j} |a_{i,j}|^2$. For the matrix 2-norm the condition number of a square nonsingular matrix is defined as $\kappa(A) = \|A\|\,\|A^{-1}\|$, and analogously for the Frobenius norm. The notation $[A; B]$ will be often used to express the matrix obtained by stacking the matrix B below the matrix A, both having conforming dimensions.

For given matrices $A \in \mathbb{C}^{n_A \times m_A}$, $A = (a_{ij})_{i=1,\ldots,n_A,j=1,\ldots,m_A}$, and $B \in \mathbb{C}^{n_B \times m_B}$, the Kronecker product is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1m_A}B \\ a_{21}B & a_{22}B & \cdots & a_{2m_A}B \\ \vdots & & & \vdots \\ a_{n_A1}B & a_{n_A2}B & \cdots & a_{n_Am_A}B \end{bmatrix} \in \mathbb{C}^{n_An_B \times m_Am_B}; \tag{3}$$

the vec operator stacks the columns of a matrix $X = [x_1, \ldots, x_m] \in \mathbb{C}^{n \times m}$ one after another as

$$\mathrm{vec}(X) = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \in \mathbb{C}^{nm \times 1}.$$

We summarize some well-known properties of the Kronecker product in the following lemma; see, e.g., [131].

Lemma 1. *Some properties:*

(i) *$vec(AXB) = (B^\top \otimes A)vec(X)$;*[2]

(ii) *If $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{m \times m}$, and $\lambda_A \in \mathrm{spec}(A)$, $\lambda_B \in \mathrm{spec}(B)$, then $\lambda_A\lambda_B \in \mathrm{spec}(A \otimes B)$ (and every eigenvalue of $A \otimes B$ is the product of eigenvalues of A and B);*

(iii) *With the notation of* (ii)*, $\lambda_A + \lambda_B \in \mathrm{spec}(I_m \otimes A + B \otimes I_n)$ (and every eigenvalue of $I_m \otimes A + B \otimes I_n$ is the sum of eigenvalues of A and B).*

3. Applications.

Matrix equations are ubiquitous in signal processing, control, and systems theory; see, e.g., [4], [247], [90], [64], [32], [218], [60] and references therein. Most time-dependent models accounting for the prediction, simulation, and control of real world phenomena may be represented as linear or nonlinear dynamical systems. Therefore, the relevance of matrix equations within engineering applications largely explains the great effort put forth by the scientific community into their numerical solution.

Linear matrix equations have an important role in the stability analysis of linear dynamical systems and also take part in the theoretical developments of nonlinear

[2]Note the transposition without conjugation for B. For real matrices we shall not make this distinction in what follows.

systems. Consider the continuous-time linear system[3]

$$\dot{x} = Ax + B_1 u, \qquad y = B_2^* x, \tag{4}$$

where x is the model state, u is the input, y is the output, and the matrices A, B_1, and B_2 are time-invariant. Assuming A is stable, that is, its eigenvalues have negative real part, then the solutions $\mathbf{P}$ and $\mathbf{Q}$ to the Lyapunov equations

$$A\mathbf{P} + \mathbf{P}A^* + B_1 B_1^* = 0, \quad A^*\mathbf{Q} + \mathbf{Q}A + B_2 B_2^* = 0$$

are called the controllability and observability Gramians, respectively, and they are used, for instance, to measure the energy transfers in the system (4); see [4, sec. 4.3.1]. Under certain additional hypotheses it may be shown that the symmetric matrices $\mathbf{P}$ and $\mathbf{Q}$ are positive definite. These latter two matrices are key when one is interested in reducing the original system into one of much smaller dimension, while essentially preserving the main dynamical system properties. Indeed, *balanced reduction*, which was originally used to improve the sensitivity to round-off propagation in filter design [188], determines an appropriate representation basis for the system such that the Gramians are equal and diagonal [4], so that the reduction of that basis will maintain this property of the Gramians. The diagonal Gramians then contain information on the output error induced by the reduced model.

Alternatively, if B_1 and B_2 have the same number of columns, one can solve the Sylvester equation

$$A\mathbf{W} + \mathbf{W}A + B_1 B_2^* = 0,$$

thus obtaining the *cross-Gramian* $\mathbf{W}$ [86], which contains information on controllability and observability of the system. For B_1, B_2 with a single column, or for A symmetric and B_1, B_2 such that $B_2^*(zI - A)^{-1}B_1$ is symmetric, it is possible to show that $\mathbf{W}^2 = \mathbf{PQ}$, so that the eigenvalues of $\mathbf{W}$ coincide with the square root of the eigenvalues of $\mathbf{PQ}$ [87], [234]. In general, the latter are called the Hankel singular values of the system, and they satisfy important invariance properties; see [4] for a detailed discussion of these quantities and their role in model order reduction. A different Sylvester equation was used in [91] to derive a numerical algorithm that couples the two Gramians $\mathbf{P}$ and $\mathbf{Q}$. Similar results can be stated for the case of the discrete-time time-invariant linear systems

$$\begin{aligned} x(k+1) &= Ax(k) + B_1 u(k), \\ y(k) &= B_2^* x(k), \end{aligned} \tag{5}$$

which are associated, for instance, with the discrete-time Lyapunov equation

$$A\mathbf{X}A^* - \mathbf{X} + B_1 B_1^* = 0.$$

As a particular case of the linear equation in (1), the generalized Lyapunov equation

$$A\mathbf{X}E^* + E\mathbf{X}A^* = C \tag{6}$$

has a special interest in control; see also recent applications in Hopf bifurcation iden-

[3] In the control literature, B_1, B_2 are usually denoted by B and C^*, respectively; we opted for a slightly different notation because here B and C have a different meaning.

tification in linear stability analysis [82], [184]. The case $E \neq I$ may arise in a control problem, for instance, when a second- or higher-order ordinary differential equation is discretized. Consider the linear time-invariant second-order system

$$\begin{aligned} Mq''(t) + Dq'(t) + Kq(t) &= B_2 u(t), \\ C_2 q'(t) + C_1 q(t) &= y(t), \end{aligned}$$

where $q(t) \in \mathbb{R}^n$ is the displacement and $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^p$ are the control input and output, respectively. Then, by defining the matrices

$$E = \begin{bmatrix} I & \\ & M \end{bmatrix}, \quad A = \begin{bmatrix} 0 & I \\ -K & -D \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ B_2 \end{bmatrix}, \quad \text{and} \quad C = [C_1, C_2],$$

the second-order system can be rewritten as a first-order linear system

$$Ex'(t) = Ax(t) + Bu(t), \qquad y(t) = Cx(t),$$

with $x(t) = [q(t); q'(t)]$, whose stability analysis gives rise to (6).

The Sylvester equation is classically employed for the design of Luenberger observers [182]; we refer the reader to section 7 for a more detailed discussion. Linear matrix equations are also used in control theory as a technical tool for solving other problems (see, e.g., [270], [90], [170], [193]) and for the reduction of nonlinear models; see, e.g., [218], [32] and references therein.

The Sylvester equation often occurs in linear and generalized eigenvalue problems for the computation of invariant subspaces by means of the Riccati equation [238], [229], [68]. In fact, the algebraic Riccati equation itself, defined in the symmetric case as

$$A^*\mathbf{X} + \mathbf{X}A - \mathbf{X}F\mathbf{X} + G = 0 \tag{7}$$

with F and G symmetric, provides a formidable setting for linear matrix equations: this quadratic equation is sometimes dealt with by solving a sequence of linear Sylvester or Lyapunov equations with possibly varying known term and coefficient matrices. The following Newton–Kleinman iteration is one of the leading methods for solving (7) in the large scale case, whenever $F = BB^*$ and $G = C^*C$ have low rank:

Algorithm 1. Given $\mathbf{X}_0 \in \mathbb{R}^{n\times n}$ such that $\mathbf{X}_0 = \mathbf{X}_0^*$, $A^* - \mathbf{X}_0 BB^*$ is stable:
1. **For** $k = 0, 1, \ldots$ until convergence:
2. **Set** $\mathcal{A}_k^* = A^* - \mathbf{X}_k BB^*$.
3. **Set** $\mathcal{C}_k^* = [\mathbf{X}_k B, \ C^*]$.
4. **Solve** $\mathcal{A}_k^* \mathbf{X}_{k+1} + \mathbf{X}_{k+1}\mathcal{A}_k + \mathcal{C}_k^*\mathcal{C}_k = 0$.

At each iteration the most computationally intensive operation is step 4, which requires the solution of a Lyapunov equation whose data changes at each iteration [42].

With the aim of controlling resonance modes in vibrating structures, Sylvester equations also arise in solving quadratic eigenvalue assignment problems; see, e.g., [48]. Large eigenvalue problems are also a key step in the detection of a Hopf bifurcation in large scale dynamical systems that depend on some physical parameters. However, it is possible to compute these parameters without actually computing the relevant eigenvalues. In [184], it was shown that this can be performed by means of a *matrix* inverse iteration procedure, which involves approximately solving a sequence of large scale Lyapunov equations; see also [82]. Lyapunov equations are also a theoretical and computational tool in the hydrodynamic stability theory of time-dependent problems, which is emerging as an attractive alternative to classical modal analysis in the quantitative description of short-term disturbance behaviors [219]. A large list of references on application problems where the Lyapunov equation plays an important role is available in the last chapter of [90].

Different application areas have emerged that can take advantage of an efficient solution of linear matrix equations. Problems associated with image processing seem to provide a rich source. For instance, Sylvester equations can be used to formulate the problem of restoration of images affected by noise [53]. The degraded image can be written as $\mathbf{g} = \mathbf{f} + \eta$, where η is the Gaussian noise vector. A linear operator (filter) L is applied to $\mathbf{g}$ to determine an estimate $\hat{\mathbf{f}} := L\mathbf{g}$ of the original image. A possible choice for L is the Wiener filter $L = \Phi_f(\Phi_f + \Phi_\eta)^{-1}$, where Φ_η is the covariance matrix of the noise, while $\Phi_f = \Phi_y \otimes \Phi_x$ is the covariance of $\mathbf{f}$, assuming that the variability in the vertical (y) and horizontal (x) directions are unrelated. The minimum mean square error estimate $\hat{\mathbf{f}}$ of $\mathbf{f}$ can be computed by solving the linear system $(I + \Phi_\eta \Phi_f^{-1})\hat{\mathbf{f}} = \mathbf{g}$. For $\Phi_\eta = \sigma_\eta^2 I$, corresponding to white and Gaussian noise η with variance σ_η^2, the system is given by

$$(I + \sigma_\eta^2 \Phi_y^{-1} \otimes \Phi_x^{-1})\hat{\mathbf{f}} = \mathbf{g},$$

which is nothing more than the Kronecker formulation of a Sylvester equation.

A similar optimization model can be used in adaptive optics, a technology developed for the compensation of aberrations in optical systems or due to atmospheric turbulence, which is mainly used in high quality astronomical observations and measurements [208]. Within the image processing application, the problem of estimating a three-dimensional object's pose obtained from two-dimensional image sequences can be stated as a constrained optimization problem [57]. This leads to the solution of a sequence of small Sylvester equations. In fact, depending on the number of poses, these linear matrix equations have more than two terms and can be formulated as in (2); see [57].

The Sylvester equation was highlighted as a *model problem* in the solution of elliptic boundary value problems governed by the two-dimensional differential operator

$$\mathcal{L}(u) = -\nabla \cdot (\kappa \nabla u)$$

by Ellner and Wachspress [80], who devised a matrix algorithmic version of the (differential) alternating-direction-implicit (ADI) algorithm of Peaceman and Rachford, and this became the foundation of ADI-type methods for linear matrix equations. Wachspress showed that the constant coefficient second-order differential equation can be used as a preconditioner for the original operator, and that the application of the preconditioner amounts to solving a Lyapunov equation [254]. Sylvester equations can also be used in the implementation of implicit Runge–Kutta integration formulae and block multistep formulae for the numerical solution of ordinary differential equations [84].

Discrete-time Sylvester and Lyapunov equations (see section 6) also arise, for instance, in statistics and probability [152], [151], [150], [10], and as a building block for solving the discrete-time algebraic Riccati equation [42].

Similarly to the Sylvester equation, the multiterm matrix equation (2) may be viewed as a model problem for certain convection-diffusion partial differential equations. For instance, let us consider the following two-dimensional problem with *separable* coefficients:

$$-\varepsilon u_{xx} - \varepsilon u_{yy} + \phi_1(x)\psi_1(y)u_x + \phi_2(x)\psi_2(y)u_y = f, \qquad (x,y) \in \Omega, \tag{8}$$

with $\varepsilon > 0$ and, for the sake of simplicity, $\Omega = (0,1) \times (0,1)$ with zero Dirichlet boundary conditions. Using standard centered finite difference discretization for each term and setting $\mathbf{U}_{ij} := u(x_i, y_j)$, where (x_i, y_j) are interior grid nodes, $i, j = 1, \ldots, n$, we obtain

$$T\mathbf{U} + \mathbf{U}T + \Phi_1 B \mathbf{U} \Psi_1^* + \Psi_2 \mathbf{U} (\Phi_2 B)^* = F, \qquad F = (f(x_i, y_j)); \tag{9}$$

here

$$T = -\frac{\varepsilon}{h^2}\ \mathrm{tridiag}(1, -2, 1), \qquad B = \frac{1}{2h}\ \mathrm{tridiag}(-1, 0, 1),$$

and

$$\Phi_k = \mathrm{diag}(\phi_k(x_1), \ldots, \phi_k(x_n)), \qquad \Psi_k = \mathrm{diag}(\psi_k(y_1), \ldots, \psi_k(y_n)), \qquad k = 1, 2,$$

where h is the mesh size. Equation (9) is a four-term linear matrix equation in $\mathbf{U}$ and was used in the early literature on difference equations; we refer the reader to, e.g., [41] for similar derivations. Common strategies then transform the problem above into the following standard real nonsymmetric linear system by means of the Kronecker product:

$$\Big(I \otimes T + T \otimes I + \Psi_1 \otimes (\Phi_1 B) + (\Phi_2 B) \otimes \Psi_2\Big)\mathbf{u} = \widetilde{f}, \qquad \mathbf{u} := \mathrm{vec}(U), \widetilde{f} = \mathrm{vec}(F),$$

for whose solution a vast literature is available. We are unaware of any recent strategies that exploit the matrix equation formulation of the problem for its numerical solution, although the matrix structure may suggest particular preconditioning strategies.

In the context of dynamical system analysis, multiterm matrix equations of the type (2) arise in the numerical treatment of bilinear systems in the form (see, e.g., [118], [217])

$$\dot{x}(t) = (A + u(t)N)x(t) + Bu(t), \quad x(0) = x_0, \quad y(t) = Cx(t), \tag{10}$$

which occur when the model accounts for a stochastic component by means of the term involving N. Other generalizations of Gramians can thus be considered, which can be written as the solution $\mathbf{X}$ to the multiterm linear matrix equation

$$A\mathbf{X} + \mathbf{X}A^* + N\mathbf{X}N^* + BB^* = 0,$$

together with its counterpart with respect to C^*C; note that extra terms of the form $N_i \mathbf{X} N_i^*$ can be included in the sum; see [23] and references therein. The solution $\mathbf{X}$ carries information on the reachability and observability properties of the state vectors [107]. The one above is an example of linear jump systems (see [183]), in which the linear coefficient matrices depend on a Markovian random process, giving rise to systems of matrix equations with an extra term, accounting for the probabilistic nature of the problem.

Another typical emerging setting where the multiterm matrix equation in (2) arises is the analysis of uncertainty quantification in data modeling. For instance, the stochastic steady state diffusion equation with homogeneous Dirichlet boundary conditions is given by

$$\begin{cases} -\nabla \cdot (c \nabla p) = f & \text{in } D \times \Omega, \\ p = 0 & \text{on } \partial D \times \Omega, \end{cases} \tag{11}$$

where D is a sufficiently regular spatial domain and Ω is a probability sample space. Both the forcing term f and the diffusion coefficient c have a stochastic component. By properly discretizing the *weak* formulation of (11), and under certain assumptions on the stochastic discretized space, one obtains the algebraic linear system (see, e.g., [85] and references therein)

$$\mathcal{A}\mathbf{p} = f, \qquad \mathcal{A} = G_0 \otimes K_0 + \sum_{r=1}^{m} \sqrt{\lambda_r} G_r \otimes K_r. \tag{12}$$

By passing to the matrix formulation and introducing the matrix $\mathbf{X}$ of coefficients in $\mathbf{p}$, (12) can be rewritten as

$$K_0 \mathbf{X} G_0^* + \sum_{r=1}^{m} \sqrt{\lambda_r} K_r \mathbf{X} G_r^* = F, \tag{13}$$

where F contains the components of f and each column of F corresponds to a different basis element in the probability space. In many simulations, while the underlying mathematical formulation is still (11), the quantity of interest is $c\nabla p$, rather than p. Using, for instance, the derivation in [89], a direct approximation to $c\nabla p$ is obtained by introducing the variable (flux) $\vec{u} = c\nabla p$, which gives

$$\begin{cases} c^{-1}\vec{u} - \nabla p = 0 & \text{in } D \times \Omega, \\ -\nabla \cdot \vec{u} = f & \text{in } D \times \Omega, \end{cases} \qquad p = 0 \quad \text{on} \quad \partial D \times \Omega. \tag{14}$$

By means of a discretization with proper (tensor products of) finite element spaces of the weak formulation of (14) (see, e.g., [89], [85], [203]), one obtains the following saddle point algebraic linear system:

$$\begin{bmatrix} A & B^* \\ B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} 0 \\ f \end{bmatrix}, \quad A = G_0 \otimes K_0 + \sum_{r=1}^{m} \sqrt{\lambda_r} G_r \otimes K_r, \ B = G_0 \otimes B_0. \tag{15}$$

The solution vectors $\mathbf{u}$ and $\mathbf{p}$ contain the two-dimensional coefficients of the (discrete) expansions of $\vec{u}$ and p column by column. Once again, a closer look at the two equations above reveals that the matrix formulation could replace the Kronecker products. Indeed, if $\mathbf{U}$ is the matrix such that $\mathbf{u} = \mathrm{vec}(\mathbf{U})$, whose coefficients are $(u_{j\ell})$, and similarly for $\mathbf{P}$, then the linear system above reads

$$K_0 \mathbf{U} G_0^* + \sum_{r=1}^{m} \sqrt{\lambda_r} K_r \mathbf{U} G_r^* + B_0^* \mathbf{P} G_0 = 0, \tag{16}$$

$$B_0 \mathbf{U} G_0^* = F, \tag{17}$$

with obvious meaning for F. This system is a natural generalization of the case in (13) and may be thought of as a saddle point *generalized matrix* system. Such systems of linear matrix equations will be discussed in section 7.2.

4. Continuous-Time Sylvester Equation.

The continuous-time Sylvester equation is possibly the most broadly employed linear matrix equation and is given as

$$A\mathbf{X} + \mathbf{X}B = C, \tag{18}$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times m}$, and $C \in \mathbb{R}^{n\times m}$. 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.

A general result on the consistency of the Sylvester equation was given by Roth in 1952 in [211] and reads: *equation* (18) *admits a solution if and only if the matrices*

$$\begin{bmatrix} A & -C \\ 0 & -B \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix} \tag{19}$$

are similar; the similarity transformation matrix is given by

$$\begin{bmatrix} I & \mathbf{X} \\ 0 & I \end{bmatrix},$$

where $\mathbf{X}$ is the solution to (18).

Using the Kronecker product, the matrix equation in (18) can be rewritten as the standard (vector) linear system

$$\mathcal{A}\mathbf{x} = c, \qquad \text{with} \quad \begin{array}{l} \mathcal{A} = I_m \otimes A + B^* \otimes I_n, \\ \mathbf{x} = \mathrm{vec}(\mathbf{X}), \quad c = \mathrm{vec}(C), \end{array} \tag{20}$$

from which we can deduce that the system admits a solution for any c, and this is unique if and only if the matrix $\mathcal{A}$ is nonsingular. Taking into account Lemma 1(iii), this is equivalent to requiring that $\mathrm{spec}(A) \cap \mathrm{spec}(-B) = \emptyset$ (see, e.g., [131, Thm. 4.4.6]). In what follows we shall thus always assume that this latter condition is satisfied, so that the solution to (18) exists and is unique; standard matrix analysis books describe the case when this spectral condition is not satisfied (see, e.g., [131], [168]). The homogeneous case, namely, when $C = 0$, can be handled correspondingly: the matrix equation has only the trivial solution $\mathbf{X} = 0$ if and only if $\mathrm{spec}(A) \cap \mathrm{spec}(-B) = \emptyset$ [97, sec. 17.8].

The solution $\mathbf{X}$ of (18) may be written in closed form in a number of different ways. These forms were derived in different references throughout the 1950s and 1960s, with contributions by E. Heinz, A. Jameson, M. G. Krein, E. C. Ma, M. Rosenblum, and W. E. Roth, among others. A beautiful account of these early contributions can be found in the survey by P. Lancaster [165], to which we also refer the reader for the bibliographic references. Here we report the main closed forms:

(a) *Integral of resolvents.* The following representation, due to Krein, exploits spectral theory arguments:

$$\mathbf{X} = -\frac{1}{4\pi^2} \int_{\Gamma_1} \int_{\Gamma_2} \frac{(\lambda I_n - A)^{-1} C (\mu I_m - B)^{-1}}{\lambda + \mu} d\mu d\lambda, \tag{21}$$

where Γ_1, Γ_2 are contours containing and sufficiently close to the spectra of A and B, respectively.

(b) *Integral of exponentials.* This representation, due to Heinz, is tightly connected to the previous one:

$$\mathbf{X} = -\int_0^\infty e^{At} C e^{Bt} dt, \tag{22}$$

where e^{Ht} is the matrix exponential of Ht. Here the spectra of A and B are assumed to be separated by a vertical line.

(c) *Finite power sum.* Let $C = C_A C_B^*$. Let a_m of degree m be the minimal polynomial of A with respect to C_A, namely, the smallest degree monic polynomial such that $a_m(A)C_A = 0$. Analogously, let b_k of degree k be the minimal polynomial of B with respect to C_B. Then

$$\begin{aligned}\mathbf{X} &= \sum_{i=0}^{m-1}\sum_{j=0}^{k-1}\gamma_{ij}A^iCB^j \\ &= [C_A, AC_A, \ldots, A^{m-1}C_A](\gamma \otimes I)\begin{bmatrix} C_B^* \\ C_B^*B \\ \vdots \\ C_B^*B^{k-1}\end{bmatrix},\end{aligned} \tag{23}$$

where γ is the solution of the Sylvester equation with coefficient matrices given by the companion matrices of a_m and b_k and right-hand side given by the matrix $[1; 0; \ldots; 0][1, 0, \ldots, 0]$ [69]; a block version of this result using minimal *matrix* polynomials can also be derived [225].

(d) *Similarity transformations.* Strictly related to (c), in addition this form assumes that A and B can be diagonalized, $U^{-1}AU = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $V^{-1}BV = \mathrm{diag}(\mu_1, \ldots, \mu_m)$. Let $\widetilde{C} = U^{-1}CV$. Then

$$\mathbf{X} = U\widetilde{X}V^{-1}, \quad \text{with} \quad \widetilde{\mathbf{x}}_{ij} = \frac{\widetilde{c}_{ij}}{\lambda_i + \mu_j}.$$

Other representations can be found in [165] and, for more general equations of the same type, in [262]. We also mention that the columns of $[\mathbf{X}; I]$ span an invariant subspace for the left matrix in (19), that is,

$$\begin{bmatrix} A & -C \\ 0 & -B \end{bmatrix}\begin{bmatrix}\mathbf{X} \\ I\end{bmatrix} = \begin{bmatrix}\mathbf{X} \\ I\end{bmatrix}S, \tag{24}$$

where it holds that $S = -B$. Equation (24) has been used both to derive matrix properties of the solution $\mathbf{X}$ and also to construct solution devices.

In [69] the closed form in (c) is used to derive results on the solution rank; results on the nonsingularity of the solution based on the same conditions are also given in [119]. For more general equations, corresponding nonsingularity conditions can be found, e.g., in [261]. In [69], the controllability (resp., observability) of the pair (A, C_A) (resp., (B^*, C_B)) plays a crucial role.[4]

Early computational methods relied on one of the analytic expressions above; see the account of early computational methods in [90]. Although these closed forms are no longer used to solve the Sylvester equation numerically, they have motivated several successful methods and they represent an important starting point for theoretical investigations of numerical approaches.

4.1. Stability and Sensitivity Issues of the Sylvester Equation. In this section we provide a brief account of the sensitivity issues encountered when solving the Sylvester equation. The topic is broad, and it also involves the solution of related matrix equations; we refer to the thorough treatment in [157] for a full account of the perturbation theory of this and other important equations in control.

The sensitivity to perturbations of the solution $\mathbf{X}$ to (18) is inversely proportional to the separation between A and $-B$, where the separation function of two matrices A_1 and A_2 is defined as

$$\mathrm{sep}_p(A_1, A_2) = \min_{\|P\|_p=1}\|A_1P - PA_2\|_p,$$

[4] A pair (M, C) is controllable if the matrix $[C, MC, \ldots, M^{n-1}C]$ has full row rank n, equal to the row dimension of M; (M, C^*) is observable if (M^*, C) is controllable.

with $p = 2, F$; see, e.g., [238]. This can be seen by recalling that the columns of $[\mathbf{X}; I]$ are a basis for an invariant subspace for the first block matrix in (24). We refer the reader to, e.g., [98, sec. 7.6.3], where the role of $\|\mathbf{X}\|_F$ in the conditioning of the associated eigenvalues is emphasized. More specifically, it holds that

$$\|\mathbf{X}\|_F \leq \frac{\|C\|_F}{\mathrm{sep}_F(A, -B)}. \tag{25}$$

For nonnormal matrices, the bound above suggests that a good spectral distance between A and $-B$ might not be sufficient to limit the size of $\|\mathbf{X}\|_F$, since $\mathrm{sep}_F(A, -B)$ can be much smaller than the distance between the spectra of A and $-B$. The function sep plays the role of a condition number for the Sylvester operator

$$\mathcal{S} : \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}, \quad \mathcal{S}(X) = AX + XB; \tag{26}$$

numerical estimates for the sep function can be obtained by carefully adapting classical strategies [49]. The occurrence of the sep function in the bound (25) suggests that for small scale equations, algorithms that rely on orthogonal reduction should be preferred for numerical stability. Methods that rely on more general transformations $\widetilde{\mathbf{X}} = U\mathbf{X}V^{-1}$ may transfer the ill-conditioning of the transformation matrices U and V onto large errors in the obtained solution; moreover (see, e.g., [238, Exercise V.2.1]),

$$\frac{\mathrm{sep}(A, B)}{\kappa(U)\kappa(V)} \leq \mathrm{sep}(UAU^{-1}, VBV^{-1}) \leq \kappa(U)\kappa(V)\mathrm{sep}(A, B).$$

A major difference between matrix equations and standard linear systems lies in their stability properties. In particular, a small Sylvester equation residual does not necessarily imply a small backward error [124, sec. 15.2]. Define the backward error for an approximation $\mathbf{X}$ as

$$\begin{aligned} \eta(\mathbf{X}) := \min\{\varepsilon : (A + \Delta A)\mathbf{X} + \mathbf{X}(B + \Delta B) = C + \Delta C, \\ \|\Delta A\|_F \leq \varepsilon\|A\|_F, \|\Delta B\|_F \leq \varepsilon\|B\|_F, \|\Delta C\|_F \leq \varepsilon\|C\|_F\}, \end{aligned}$$

and the residual as $R = C - (A\mathbf{X} + \mathbf{X}B)$. Then [123]

$$\eta(\mathbf{X}) \leq \mu \frac{\|R\|_F}{(\|A\|_F + \|B\|_F)\|\mathbf{X}\|_F + \|C\|_F}, \tag{27}$$

where μ is an amplification factor depending on the data norms and on the singular values of $\mathbf{X}$. For instance, for $n = m$ this factor has the expression

$$\mu = \frac{(\|A\|_F + \|B\|_F)\|\mathbf{X}\|_F + \|C\|_F}{((\|A\|_F^2 + \|B\|_F^2)\sigma_{\min}(\mathbf{X})^2 + \|C\|_F^2)^{\frac{1}{2}}},$$

making the dependence on the norm and ill-conditioning of $\mathbf{X}$ more apparent. A more complex situation occurs for $n \neq m$; we refer the reader to [124, sec. 15.2] for more details, and to [157] for a more thorough perturbation analysis. We also mention that in [243] bounds for the norm of the solution $\mathbf{X}$ and of its perturbation are obtained that emphasize the influence of the possibly low-rank right-hand side on the sensitivity of the solution itself. The distribution of the singular values of $\mathbf{X}$ plays a crucial role in the stability analysis of dynamical systems and also in the quality of low-rank approximations. In section 4.4 we recall some available estimates for the singular values that also motivate the development of low-rank approximation methods.

4.2. Sylvester Equation. Small Scale Computation. A robust and efficient method for numerically solving Sylvester equations of small and moderate size was introduced in 1972 by Bartels and Stewart [15], and with some modifications is still the state of the art; in section 8 we give an account of current software, much of which relies on this method. The idea is to compute the Schur decomposition of the two coefficient matrices and then transform the given equation into an equivalent one that uses the quasi-lower/upper triangular structure of the Schur matrices. This last equation can then be explicitly solved element by element. To introduce the algorithm, let us first consider the general case of complex A and B. Then the following steps are performed (see, e.g., [98]):

Algorithm 2.
1. Compute the Schur forms $A^* = URU^*$, $B = VSV^*$ with R, S upper triangular.
2. Solve $R^*\mathbf{Y} + \mathbf{Y}S = U^*CV$ for $\mathbf{Y}$.
3. Compute $\mathbf{X} = U\mathbf{Y}V^*$.

The Schur forms in the first step are obtained by the QR iteration [98], while the third step is a simple product. It remains to explain how to solve the new structured Sylvester equation in the second step. Since R^* is lower triangular and S is upper triangular, the (1,1) element of $\mathbf{Y}$ can be readily obtained. From there the next elements of the first row in $\mathbf{Y}$ can also be obtained sequentially. Similar reasoning can be used for the subsequent rows.

In the case of real A and B, the *real* Schur form may be exploited, where R and S are now quasi-triangular, that is, the diagonals have 2×2 and 1×1 blocks corresponding to complex and real eigenvalues, respectively. The solution process can then use the equivalence between a 2×2 Sylvester equation and the associated Kronecker form in (20); see, e.g., [223, sec. 2.3.1]. The same sequential process as in the complex case can be employed to compute the elements of $\mathbf{Y}$, as long as the diagonal blocks can be made conforming; for nonconforming dimensions, a sequence of small shifted linear systems needs to be solved; the details can be found in [15], [98, sec. 7.6.3], [223, sec. 2.3.1]. The method outlined above is at the core of most linear matrix equation solvers in software packages such as LAPACK[5] and SLICOT [246], [232], [27]. The leading computational cost is given by the Schur forms in the first step, which for real matrices are nowadays performed in real arithmetic. Explicitly computing the Schur form costs at least $10n^3$ floating point operations for a matrix of size n [98]; to limit costs, the Bartels–Stewart algorithm is commonly employed only if either A or B is already in Schur or upper Hessenberg form; see, e.g., [232]. For general matrices A and B, the method proposed by Golub, Nash, and Van Loan in 1979 [99] can be considerably faster, especially if either m or n is significantly smaller than the other. This latter method replaces the Schur decomposition of the larger matrix, say, B, with the Hessenberg decomposition of the same matrix whose computational cost is $5/3m^3$, which should be compared with $10m^3$ for the Schur form [99]. We refer the reader to [223, sec. 2.3.1] for a more detailed comparison of the computational costs. In [236], a variant of the Bartels–Stewart algorithm is proposed: the forward-backward substitution in step 2 is performed by a columnwise block scheme, which seems to be better suited for modern computer architectures than the original complex version. In [143], [144], the authors propose an even more effective implementation based on splitting the matrices, already in block triangular

[5] http://www.netlib.org/lapack/.

form, and then recursively solving for each block. For instance, if A is much larger than B ($n \geq 2m$), then the original equation can be written as

$$\begin{bmatrix} A_{11} & A_{12} \\ & A_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} B = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix},$$

with obvious meaning for the blocks. The second block equation gives the smaller size Sylvester equation $A_{22}X_2 + X_2B = C_2$, which can again be split by using the block triangular form of A_{22}, and the solution is obtained in a recursive manner. Once X_2 is fully recovered, X_1 can be computed by recursively solving with the updated right-hand side in the first block equation above. Different size cases and different triangular structures can be handled and are described in [143]. These advanced strategies have been included in the software package RECSY[6] and in LAPACK; see section 8.

Iterative solution strategies for small size matrices have also been proposed: given an initial guess $\mathbf{X}_0$, they determine a sequence of matrices $\mathbf{X}_1, \ldots, \mathbf{X}_k, \ldots$ that converge to $\mathbf{X}$. These are related to a basic Newton iteration for approximating the matrix sign function. In section 5.2.3 we will give more details in relation to the Lyapunov equation, although the procedure can be used for stable Sylvester equations as well [36]. These approaches are easier to parallelize than QR-based methods. For instance, it is shown in [36] that they provide high efficiency and scalability on clusters of processors.

To conclude, a special mention should be made of the Sylvester equation with $B = -A$, yielding the so-called *displacement equation*

$$A\mathbf{X} - \mathbf{X}A = C, \tag{28}$$

which measures how far A and $\mathbf{X}$ are from commuting; see, e.g., [96] for typical applications in the context of structured matrices such as Cauchy-like and Toeplitz matrices.

4.3. Sylvester Equation. Large *A* and Small *B*. When either n or m is large, Schur factorization may require a prohibitive amount of space, due to the dense nature of the corresponding large matrix. Selecting the most appropriate solver still depends on whether the smaller matrix has very small dimension. Different approaches can then be used when decomposing the small matrix is feasible.[7] To fix ideas, and without loss of generality, we shall assume that B is small (size less than 1000) and A is large (size much bigger than 1000), so that $m \ll n$.

In this section we thus consider that the equation can be visualized as

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} \\ \mathbf{X} \\ \\ \end{bmatrix} + \begin{bmatrix} \\ \mathbf{X} \\ \\ \end{bmatrix} \begin{bmatrix} B \end{bmatrix} = \begin{bmatrix} \\ C \\ \\ \end{bmatrix}, \tag{29}$$

so that the large dimension of A makes the methods discussed in section 4.2 unfeasible. This situation arises, for instance, in the solution of eigenvalue problems [258, sec. 2.4, sec. 6.6] and in (separable) boundary value problems [254], [256], [41]. We immediately notice that for very small m, the transformation with the Kronecker product (20) might be appealing, since the dimension of the linear system might be just a few (m)

[6]http://www8.cs.umu.se/~isak/recsy/.

[7]Feasibility is machine architecture dependent; nonetheless, a matrix of dimension much less than 1000 should be considered small.

times that of A. However, projection methods acting on the original matrix equation turn out to be extremely effective in this case, possibly explaining the sparsity of attempts to pursue the Kronecker formulation. We next describe some of the standard approaches currently employed in the literature and in applications.

Assume that B can be spectrally decomposed cheaply and stably. Then by writing $B = WSW^{-1}$ with $S = \text{diag}(s_1, \ldots, s_m)$, we obtain

$$A\widehat{\mathbf{X}} + \widehat{\mathbf{X}}S = \widehat{C}, \quad \widehat{\mathbf{X}} = \mathbf{X}W, \quad \widehat{C} = CW. \tag{30}$$

For B Hermitian, $W^{-1} = W^*$. Each column of $\widehat{\mathbf{X}}$ can be obtained by solving a shifted linear system $(A + s_i I)(\widehat{\mathbf{X}})_i = (\widehat{C})_i$, where $(\widehat{\mathbf{X}})_i$ denotes the ith column of $\widehat{\mathbf{X}}$. The main steps can be summarized in the following algorithm:

Algorithm 3.
1. Compute the decomposition $B = WSW^{-1}$.
2. Set $\widehat{C} = CW$.
3. For $i = 1, \ldots, m$ solve $(A + s_i I)(\widehat{\mathbf{X}})_i = (\widehat{C})_i$.
4. Compute $\mathbf{X} = \widehat{\mathbf{X}}W^{-1}$.

The shifted systems in step 3 can be solved simultaneously by using standard solvers for algebraic linear systems, either direct or iterative; see, e.g., [214], [230] and their references. We also note that step 3 is "embarrassingly parallel" when different systems can be distributed on a multiprocessor machine.

If the eigendecomposition of B is not appealing, then one can resort to a (complex) Schur decomposition $B = QR_BQ^*$, giving $A\mathbf{X}Q + \mathbf{X}QR_B = CQ$. Since R_B is upper triangular, these systems can still be solved using the shifted form, but this time in sequence: letting r_{ij} be the (i, j) entry of R_B and $\widehat{C} = CQ$, we have

$$\text{for } i = 1, \ldots, m, \quad (A + r_{ii}I)(\widehat{\mathbf{X}})_i = (\widehat{C})_i - \sum_{k=1}^{i-1} r_{ki}(\widehat{\mathbf{X}})_k, \quad \widehat{\mathbf{X}} = \mathbf{X}Q. \tag{31}$$

Such an approach has been used in different contexts; see, e.g., [110], [234], [26], where the Sylvester equation considered is occasionally called a *sparse-dense* equation.

For moderate n, the use of direct methods in (30) and (31) may entail the use of complex arithmetic if the shifts s_i (eigenvalues) are complex, significantly increasing the computational cost; the alternative of solving two real systems also leads to higher computational costs. In addition, when the use of sparse direct methods appears to be competitive, it should be noted that only the sparsity analysis step can be done once, whereas the actual decomposition needs to be performed again for each distinct shift.

Major computational savings may be obtained if C is low rank, namely, $C = C_0R$, with $C_0 \in \mathbb{R}^{n \times \bar{m}}$ and $\bar{m} < m$. Indeed, the m shifted systems can be solved more efficiently by working only with the common matrix C_0. For the rest of this section we assume that C is full rank and postpone the treatment of the low-rank case to later, when we discuss the occurrence of large B. Indeed, the rank of C is key in developing general projection methods, as is explained next.

Projection Methods. Let $\mathcal{V}$ be a subspace[8] of $\mathbb{C}^n$ of dimension k, and let the columns of $V_k \in \mathbb{C}^{n \times k}$ span $\mathcal{V}$. An approximate solution $\mathbf{X}_k$ with $\text{range}(\mathbf{X}_k) \subset \mathcal{V}$ is

[8]We use complex arithmetic for $\mathcal{V}$ to allow for complex spaces also for real data, which may occur when using rational Krylov subspaces with complex shifts. A careful implementation can construct a real space if conjugate shifts are used. For the sake of generality we stick to complex arithmetic for $\mathcal{V}$.

sought such that

$$R_k := A\mathbf{X}_k + \mathbf{X}_k B - C \approx 0.$$

Several options arise, depending on the choice of $\mathcal{V}$ and the strategy to determine $\mathbf{X}_k$ within the space $\mathcal{V}$. For a given $\mathcal{V}$, thus let $\mathbf{X}_k = V_k \mathbf{Y}_k \approx \mathbf{X}$ for some $\mathbf{Y}_k \in \mathbb{C}^{k\times m}$ to be determined. Recalling the operator $\mathcal{S}$ defined in (26), we observe that $\mathcal{S}$ generalizes to the "block" B the concept of shifted matrices, namely,

$$x \mapsto (A + \beta I)x = Ax + x\beta.$$

Therefore, it is very natural to extend the algorithmic strategies of linear systems to the case of $\mathcal{S}$. Extensions of the linear system solvers CG (FOM) and MINRES (GMRES) can be thought of for A Hermitian (non-Hermitian), although the actual implementation differs. All these solvers are derived by imposing some orthogonality condition on the system residual. If we require that the columns of the matrix R_k be orthogonal to the approximation space $\mathcal{V}$ in the Euclidean inner product, then we are imposing the following Galerkin condition (see also (40)):

$$V_k^* R_k = 0 \quad \Leftrightarrow \quad (I \otimes V_k)^* \mathrm{vec}(R_k) = 0.$$

For simplicity, let us assume that $V_k^* V_k = I$. Then

$$0 = V_k^* R_k = V_k^* A V_k \mathbf{Y}_k + \mathbf{Y}_k B - V_k^* C. \tag{32}$$

The condition thus gives a new Sylvester equation of reduced size. Under the hypothesis that $\mathrm{spec}(V_k^* A V_k) \cap \mathrm{spec}(-B) = \emptyset$, (32) can be solved efficiently by one of the methods discussed in section 4.2. The procedure above holds for any space $\mathcal{V}$ and associated full-rank matrix V_k. Therefore, the effectiveness of the approximation process depends on the actual selection of $\mathcal{V}$. A well-exercised choice is given by the block Krylov subspace

$$K_k^{\square}(A, C) = \mathrm{range}([C, AC, \ldots, A^{k-1}C]). \tag{33}$$

The following result proved in [209, Lem. 2.1], [225] generalizes the well-known shift invariance property of vector Krylov subspaces to the case of blocks, where the $m \times m$ matrix B plays the role of the shift; the operator $\mathcal{S}$ is as defined in (26).

PROPOSITION 2. *Define $\mathcal{S}^j(C) = \mathcal{S}(\mathcal{S}^{j-1}(C))$, $j > 0$, and $\mathcal{S}^0(C) = C$. Then*

$$K_k^{\square}(A, C) = K_k^{\square}(\mathcal{S}, C) := \mathrm{range}([C, \mathcal{S}(C), \ldots, \mathcal{S}^{k-1}(C)]).$$

For the space in (33), the procedure outlined above is the complete analogue of that giving rise to the full orthogonalization method (FOM) for $m = 1$ or for $B = 0$. However, due to possible loss of rank in the basis, it was suggested in [209] to generate the subspace with A rather than with $\mathcal{S}$. As an example, Algorithm 4 describes an implementation of the projection method with the generation of the block Krylov subspace and the determination of the approximation by imposing the Galerkin orthogonality condition.

Algorithm 4. Given A, B, C:
1. Orthogonalize the columns of C to find $v_1 = V_1$.
2. $k = 1, 2, \ldots$
3. Compute $\mathbf{Y}_k$, solution to $(V_k^* A V_k)\mathbf{Y} + \mathbf{Y}B - V_k^* C = 0$.

4. If converged, $\mathbf{X}_k = V_k\mathbf{Y}_k$ and stop.
5. Arnoldi procedure for the next basis block:
 $\hat{v} = Av_k$.
 Make $\hat{v}$ orthogonal wrt $\{v_1, \ldots, v_k\}$.
 Orthogonalize (wrt 2-norm) the columns of $\hat{v}$ to get v_{k+1}.
 Update: $V_{k+1} = [V_k, v_{k+1}]$.

For future reference, we remark that the Arnoldi procedure used in Algorithm 4 generates a matrix recurrence that can be written as

$$AV_k = V_k H_k + \hat{v}e_k^*, \tag{34}$$

where $\hat{v}$ is the new block of basis vectors, prior to orthogonalization, and H_k contains the orthogonality coefficients with $H_k = V_k^* AV_k$.

One could consider constraint spaces different from the approximation spaces; in this case, a so-called Petrov–Galerkin condition is imposed on the residual. To this end, let us consider the matrix inner product defined as

$$\langle Y, X\rangle_F = \mathrm{trace}(Y^*X), \qquad X, Y \in \mathbb{R}^{n\times m}. \tag{35}$$

Following the standard linear system case with $m = 1$ and using, e.g., the space spanned by the columns of AV_k, one might be tempted to impose the condition $(AV_k)^*R_k = 0$ in the Euclidean inner product, giving

$$V_k^*A^*AV_k\mathbf{Y}_k + V_k^*A^*V_k\mathbf{Y}_kB - V_k^*A^*C = 0. \tag{36}$$

In the standard ($B = 0$) linear system setting, this condition is equivalent to minimizing the residual R_k in the Frobenius norm, that is,

$$\min_{\mathbf{Y}_k\in\mathbb{R}^{k\times m}} \|R_k\|_F. \tag{37}$$

However, for $B \neq 0$, such equivalence does *not* hold, that is, the solution to (36) is not a residual minimizing approximation. To attain a residual minimization, the orthogonality condition should be applied to the operator $\mathcal{S}$ in (26) in the Frobenius inner product (35); to this end, we note that the adjoint operator $\mathcal{S}^*$ with respect to the inner product in (35) is given by $\mathcal{S}^*(X) = A^*X + XB^*$.

PROPOSITION 3 (see [209, sec. 3]). *Let $\mathbf{Y}_k \in \mathbb{R}^{k\times m}$ and let $R_k = AV_k\mathbf{Y}_k + V_k\mathbf{Y}_kB - C$ be the associated residual. Then*

$$\mathbf{Y}_k = \arg\min_{\mathbf{Y}_k\in\mathbb{R}^{k\times m}} \|R_k\|_F \quad \textit{if and only if} \quad R_k \perp_F \mathcal{S}(K_m(\mathcal{S}, V_1)).$$

For the choice $\mathcal{V} = K_k^{\square}(A, C)$, the minimization process in (37) is the matrix analogue of GMRES (for $m = 1$ or $B = 0$) (see [214, sec. 6.12]). Similar results are discussed independently in [110]. Inspired by the “block shift” invariance of Proposition 2, the authors of [209] provide a detailed description of the parallels that can be drawn between solving (29) for $m \ll n$ with Galerkin and residual minimizing procedures and solving linear systems $A\mathbf{X} = C$ by means of block methods. Upper bounds for the residual norm of Galerkin and residual minimizing methods with $\mathcal{V} = K_k^{\square}(A, C)$ are also provided in [209], together with numerical experiments on the performance of the approaches.

Preconditioned global Krylov subspaces have also been proposed as approximation spaces [46], which, however, simply amount to a convenient implementation of a subspace method for the Kronecker formulation of the problem; see also section 4.4.1.

An alternative choice of approximation space $\mathcal{V}$ has recently shown great potential compared with the block Krylov subspace and is given by the *extended* Krylov subspace, defined as

$$\mathbf{EK}_k(A, C) := K_k^{\square}(A, C) + K_k^{\square}(A^{-1}, A^{-1}C). \tag{38}$$

Since the spaces are nested, namely, $\mathbf{EK}_k(A, C) \subseteq \mathbf{EK}_{k+1}(A, C)$, the space can be generated iteratively, allowing one to improve the approximate solution as the recurrence proceeds. For A large and sparse, experiments in [227] show that the good performance of the derived method seems to fully compensate for the high costs of solving linear systems with A at each iteration.

4.4. Sylvester Equation. Large *A* and Large *B*. In the most general case, both A and B have large dimensions. This setting arises in many situations, such as in the discretization of separable PDEs [80] or in the computation of the *cross*-Gramian in control [4]. A particularly important observation is that the dimensions of A and B determine that of $\mathbf{X}$, and that although A and B may be sparse, $\mathbf{X}$ is dense, in general. In this context, the distribution of the singular values of $\mathbf{X}$ plays a key role in the development and convergence analysis of iterative solution methods. Indeed, a Sylvester equation having solution with exponentially decaying singular values can be well approximated by a low-rank matrix. The possibility of writing $C = C_1 C_2^*$ with C_1, C_2 with low column rank is crucial to obtaining good low-rank approximations to $\mathbf{X}$, thus avoiding the storage of the whole matrix, which is in general prohibitive. We recall here the result described by Sabino in [215, Thm. 2.1.1], while Sabino's Ph.D. thesis contains further discussion related to this bound. Here K and K' are the complete elliptic integrals of the first kind[9] [1]. Additional considerations and results are postponed to the Lyapunov equation case considered in section 5.

THEOREM 4. *Let A and B be stable and real symmetric, with spectra contained in $[a, b]$ and $[c, d]$, respectively. Define $\eta = 2(b-a)(d-c)/((a+c)(b+d))$. Assume C is of rank p. Then the singular values $\sigma_1 \geq \cdots \geq \sigma_{\min\{m,n\}}$ of the solution $\mathbf{X}$ to* (18) *satisfy*

$$\frac{\sigma_{pr+1}}{\sigma_1} \leq \left(\frac{1 - \sqrt{k_r'}}{1 + \sqrt{k_r'}} \right)^2, \quad 1 \leq pr < n,$$

where $k_r' = 1/(1+\eta+\sqrt{\eta(\eta+2)})$ is the complementary elliptic modulus corresponding to the nome q^r, $q := \exp(-\pi K'/K)$.

A more accessible and practical estimate for $B = A (= A^*)$ and small condition number $\kappa(A) = \|A\| \, \|A^{-1}\|$ may be obtained as [215]

$$\frac{\sigma_{pr+1}}{\sigma_1} \lesssim 4\exp(-\pi^2 r / \log(4\kappa(A))). \tag{39}$$

A rule of thumb suggested in [215, Rule of Thumb 2.1.4] is that if $\kappa(A)$ is on the order of 10^q, then the ratios of σ_j/σ_1 decrease by a factor of 10 for every increase in j by $p(q+1)/2$.

[9] They are defined as $K = K(k) = \int_0^1 [(1-t^2)(1-kt^2)]^{-1/2} dt$ and $K' = K(1-k)$, with k being the modulus, $k = \sqrt{1-(k')^2}$, while the complementary elliptic modulus k' is given.

Easy-to-use variants of (39) in [215] are favorably compared with earlier estimates in [199]. Results for A and B nonsymmetric are scarce; nonnormality can strongly influence the solution numerical rank and the singular value decay, so that results depart significantly from the bound above. A satisfactory understanding of the singular value decay for nonnormal coefficient matrices is still lacking.

From a numerical analysis viewpoint, we notice that the main rational approximation ingredients used for results of the type above are the same as those obtained for rational space projections and ADI-type iterations (see sections 4.4.1 and 4.4.2), which also rely on minimax rational approximations; in fact, the result above is intimately related to similar estimates for ADI by Ellner and Wachspress in [80], [81].

Numerical methods in the literature have mainly proceeded in three directions: projection-type approaches (mostly based on the Krylov subspace family), matrix updating sequences (such as ADI iterations), and sparse data format recurrences. Combinations of these approaches have also been explored.

The convergence rates of the strategies in the first two categories above strongly depend on the spectral properties of the coefficient matrices (eigenvalues or field of values). For those problems with unfavorable spectral properties, for instance, with fields of values of A and $-B$ close to each other, the most efficient available methods rely on iterations that involve solving linear systems at each step, either with A or with $A + \sigma I$ for some appropriately chosen σ. For A large but very sparse, these solves can be conveniently carried out by means of direct methods. On the other hand, if the direct solution with A becomes prohibitively expensive, in terms of CPU time or memory requirements, an (inner) iterative solution of the linear systems with $A + \sigma I$ is performed at each step, giving rise to an inner-outer procedure. In this case, one usually talks about "inexact solves," unless the iterative process allows one to reach machine precision accuracy. Such a consideration noticeably influences the evaluation of the computational costs of these methods, whose performance is thus problem dependent.

Due to the important role the Lyapunov equation has in control problems, many authors have developed numerical procedures specifically for that equation, and not for the Sylvester equation, although in many cases they could be extended to the latter in a natural manner. For historical reasons, and also to avoid constant reference to the equation context, we will refer to the literature in the way the methods were originally presented. In particular, it will be apparent that the literature on Lyapunov equations is richer than that for the Sylvester equation, especially in the large scale case.

We also notice that, as a major distinction from linear vector equations, the numerical solution of matrix equations cannot directly rely on preconditioning strategies, unless the Kronecker formulation is employed. Indeed, preconditioning methods would necessarily destroy the symmetry properties of the problem, which allows one to deal with computational costs that depend on powers of n and m, but not on powers of $n \cdot m$. As an example, let us assume that a nonsingular matrix P exists[10] such that $P^{-1}A$ and $P^{-1}B^*$ have better spectral properties than the original matrices; for A, B symmetric, this requirement corresponds to a better clustering of the eigenvalues. Then we could consider applying P as follows:

$$P^{-1}A\mathbf{X}P^{-*} + P^{-1}\mathbf{X}BP^{-*} = P^{-1}CP^{-*}.$$

To be able to rewrite such an equation in terms of a single unknown matrix, one could

[10] One could also consider the existence of two matrices, one for A and one for B^*.

premultiply and postmultiply $\mathbf{X}$ by P^{-1} and P^{-*}, that is,

$$(P^{-1}AP)P^{-1}\mathbf{X}P^{-*} + P^{-1}\mathbf{X}P^{-*}(P^*BP^{-*}) = P^{-1}CP^{-*}.$$

Unfortunately, this transformation yields coefficient matrices that are similar to the original ones, thus the eigenvalues are unchanged. This simple example shows that different acceleration strategies need to be developed for the Sylvester equation; research has thus focused on constructing information-rich approximation spaces by using spectral transformations, rather than preconditioning as is done in eigenvalue computations.

4.4.1. Projection Methods. When both n and m are large, the dense solution matrix $\mathbf{X}$ of (18) cannot be stored, therefore the determination of a memory saving approximation becomes mandatory. Whenever $C = C_1C_2^*$ has low rank, the results discussed in section 4.4 suggest that a low-rank approximate solution can be determined, so that projection strategies are very appealing. Indeed, these methods compute low-rank approximations $\widetilde{\mathbf{X}} = V_k\mathbf{Y}W_j^* \approx \mathbf{X}$, with V_k and W_j that have far fewer columns than n and m, respectively, and are generalizations of the procedure seen in section 4.3.

Let $\mathcal{V}$ and $\mathcal{W}$ be two subspaces of $\mathbb{C}^n$, in principle not necessarily of the same dimension, and let the k (j) columns of V_k (of W_j) be orthonormal bases for $\mathcal{V}$ (for $\mathcal{W}$), with $k \ll n$, $j \ll m$, such that $\mathcal{V}$ is not orthogonal to range(C_1) and $\mathcal{W}$ is not orthogonal to range(C_2). We look for an approximation $\widetilde{\mathbf{X}} = V_k\mathbf{Y}W_j^* \approx \mathbf{X}$, and we let $R := C_1C_2^* - A\widetilde{\mathbf{X}} - \widetilde{\mathbf{X}}B$ be the associated residual. Then we have $\widetilde{\mathbf{x}} = \mathrm{vec}(\widetilde{\mathbf{X}}) = (W_j \otimes V_k)\mathrm{vec}(\mathbf{Y})$, where $\widetilde{\mathbf{x}}$ is an approximate solution of (20). Imposing a Galerkin (orthogonality) condition to the vector residual $c - \mathcal{A}\widetilde{\mathbf{x}}$ in (20) with respect to the space spanned by $W_j \otimes V_k$ corresponds to writing

$$(W_j \otimes V_k)^*(c - \mathcal{A}\widetilde{\mathbf{x}}) = 0 \quad \Leftrightarrow \quad V_k^*RW_j = 0. \tag{40}$$

Other conditions could be considered, such as the minimization of the residual in some norm, or the orthogonality of the residual with respect to some other space; see, e.g., [132], [130], [179].

If the columns of V_k and W_j span the spaces $K_k^{\square}(A, C_1)$ and $K_j{}^{\square}(B^*, C_2)$, respectively, as in (33), then the obtained approximate solution $\widetilde{\mathbf{X}} = V_k\mathbf{Y}W_j^*$ may also be written as

$$\widetilde{\mathbf{X}} = [C_1, AC_1, \ldots, A^{k-1}C_1]G[C_2, B^*C_2, \ldots, (B^*)^{j-1}C_2]^*$$

for some matrix G, showing that projection methods yield a *polynomial* approximation to $\mathbf{X}$, which may be viewed as particular truncations of the finite sum closed form of the solution in (23); see [225] for more details on this formulation.

Substituting the residual matrix in the equation $V_k^*RW_j = 0$ gives the following small size Sylvester equation:

$$V_k^*AV_k\mathbf{Y} + \mathbf{Y}W_j^*BW_j = V_k^*C_1(W_j^*C_2)^*. \tag{41}$$

If $V_k^*AV_k$ and $-W_j^*BW_j$ have disjoint spectra, then this equation admits a unique solution for any right-hand side. By assuming that the fields of values of A and $-B$ are disjoint, one can ensure that $V_k^*AV_k$ and $-W_j^*BW_j$ have disjoint spectra. Though restrictive, such an assumption is welcome also for stability purposes, to monitor that the solution $\mathbf{X}$ has moderate norm. A typical implementation which proceeds simultaneously with both spaces is depicted in Algorithm 5.

Algorithm 5. Given A, B, C_1, C_2:
1. Orthogonalize columns of C_1 to get $v_1 = V_1$.
2. Orthogonalize columns of C_2 to get $w_1 = W_1$.
3. For $k = 1, 2, \ldots$
4. Compute $\mathbf{Y}_k$, solution to $(V_k^* A V_k)\mathbf{Y} + \mathbf{Y}(W_k^* B W_k) - V_k^* C_1 C_2^* W_k = 0$.
5. If converged, return $V_k, \mathbf{Y}_k, W_k$ s.t. $\mathbf{X}_k = V_k \mathbf{Y}_k W_k^*$ and stop.
6. Compute next bases block:
   Compute $\hat{v}$ and $\hat{w}$ for the chosen approximate space.
   Make $\hat{v}$ orthogonal wrt $\{v_1, \ldots, v_k\}$ and $\hat{w}$ orthogonal wrt $\{w_1, \ldots, w_k\}$.
   Orthogonalize columns of $\hat{v}$ to get v_{k+1} and columns of $\hat{w}$ to get w_{k+1}.
   Update: $V_{k+1} = [V_k, v_{k+1}]$, $W_{k+1} = [W_k, w_{k+1}]$.

The process outlined in Algorithm 5 is very similar to that of Algorithm 4, the only difference being that here the space for B^* also needs to be generated. For C_1, C_2 with p columns and n and m rows, respectively (with, say, $m > n$), the computational cost at each iteration k can be summarized as follows:

(i) Solution of the projected problem: $\mathcal{O}((kp)^3)$ flops (see section 4.2).

(ii) Orthogonalization of the new basis vectors with respect to the older vectors: $\mathcal{O}(mkp^2)$.

(iii) Orthogonalization of the new block: $\mathcal{O}(mp^2)$.

We also recall that in the case when the generated basis experiences loss of rank, standard deflation procedures can be applied to remove redundant columns, ensuring a reduction in the number of columns of the current basis block in subsequent iterations. Loss of rank may occur independently of the presence of an invariant subspace of the coefficient matrix, simply due to the redundancy of some of the generated information; see [113] for a discussion in the context of linear systems with multiple right-hand sides.

The computational cost of generating the next basis vectors $\hat{v}$ and $\hat{w}$ and the quality of the approximation both depend on the choice of $\mathcal{V}$ and $\mathcal{W}$. This choice is usually based on similar arguments for each of the two spaces. We thus discuss the choice of $\mathcal{V}$, while the choice of $\mathcal{W}$ can be made analogously. In his seminal article [213], Saad proposed Krylov subspaces for determining a low-rank approximate solution to the Lyapunov equation by projection (the extension to the Sylvester equation is straightforward); the motivation was that Krylov subspaces tend to approximate well the action of the matrix exponential to a vector, so that the solution in the integral form (22) can take advantage of this property (see also section 5 for an explicit derivation). A major problem with this approach is that both bases V_k and W_j need to be stored to compute the final approximate solution. Since both matrices are full, this places a severe limitation on the maximum affordable size of the two Krylov subspaces when A and B are large. In the quest for small but more effective spaces, several alternatives have been investigated. The impressive performance results of these enriched spaces have led to a resurgence of projection-type methods for linear matrix equations. In addition to the standard Krylov subspace, we list here a couple of recently explored selections for $\mathcal{V}$ with A and C_1; similar choices can be made for $\mathcal{W}$ using B^* and C_2.

(a) Standard (block) Krylov subspace:

$$\mathcal{V} = \operatorname{range}([C_1, AC_1, A^2C_1, \ldots]).$$

(b) Rational (block) Krylov subspace:

$$\mathcal{V} = \text{range}([(A+\sigma_1 I)^{-1}C_1, (A+\sigma_2 I)^{-1}(A+\sigma_1 I)^{-1}C_1, \ldots])$$

for a specifically chosen sequence $\{\sigma_j\}$, $j = 1, 2, \ldots$, that ensures nonsingularity of the shifted matrix.

(c) Global Krylov subspace:

$$\mathcal{V} = \left\{ \sum_{i\geq 0} A^i C_1 \gamma_i,\ \gamma_i \in \mathbb{R} \right\} = \text{span}\{C_1, AC_1, A^2C_1, \ldots\},$$

where the linear combination is performed blockwise.

In all instances the least number of powers is computed so as to reach the dimension k. The subspaces listed above are somewhat related. For instance, the standard Krylov subspace can be formally obtained from the rational Krylov subspace for $\sigma_j = \infty$ for all j. Moreover, the rational block Krylov subspace also includes the special choice of fixed poles at zero and infinity, which corresponds to the *extended* Krylov subspace in (38), namely, $K_j^{\square}(A, C_1) + K_k^{\square}(A^{-1}, A^{-1}C_1)$, where j and k can in principle be different [75]. In addition, one can impose that C_1 belongs to the rational Krylov subspace with the choice $\sigma_1 = \infty$. The global Krylov subspace in (c) is a subspace of the block Krylov subspace; it was first proposed to solve linear systems with multiple right-hand sides [141], and was then adapted to the Sylvester equation in [138]. Global spaces may be viewed as simplified versions of block Krylov spaces, where the polynomial coefficients are chosen to be multiples of the identity matrix, therefore lowering the number of degrees of freedom.

The criterion for stopping the iterative procedure in Algorithm 5, and thus the approximation space expansion, is usually based on the Frobenius or 2-norm of the residual matrix $R = A\widetilde{X} + \widetilde{X}B - C_1C_2{}^*$. In general, R is dense and should not be computed explicitly if it has large dimensions. Its norm can be computed more cheaply if the generated spaces satisfy certain relations. Hence, assume that $\widehat{v}_k, \widehat{w}_j$, $\underline{H}_k$, and $\underline{K}_j$ exist such that $AV_k = [V_k, \widehat{v}_k]\underline{H}_k$ and $B^*W_j = [W_j, \widehat{w}_j]\underline{K}_j$, where $[V_k, \widehat{v}_k]$ and $[W_j, \widehat{w}_j]$ have orthonormal columns. If C_1 and C_2 satisfy $C_1 = [V_k, \widehat{v}_k]C_1^{(k)}$, $C_2 = [W_j, \widehat{w}_j]C_2^{(j)}$ for some $C_1^{(k)}, C_2^{(j)}$, then

$$\begin{aligned}\|R\|_F &= \|AV_k\mathbf{Y}W_j^* + V_k\mathbf{Y}W_j^*B - \widehat{V}_kC_1^{(k)}(\widehat{W}_jC_2^{(j)})^*\|_F \\ &= \|[V_k, \widehat{v}_k](\underline{H}_k\mathbf{Y}[I,0] + [I;0]\mathbf{Y}\underline{K}_j^* - C_1^{(k)}(C_2^{(j)})^*)[W_j, \widehat{w}_j]^*\|_F \\ &= \|\underline{H}_k\mathbf{Y}[I,0] + [I;0]\mathbf{Y}\underline{K}_j^* - C_1^{(k)}(C_2^{(j)})^*\|_F. \qquad (42)\end{aligned}$$

The last expression involves a small matrix if k and j are small, and thus its norm can be cheaply evaluated. The spaces (a) to (c) above do satisfy the required conditions, and thus the residual norm can be monitored as the iteration proceeds.

All spaces listed above are nested, so that an approximate solution can be derived while each of them is expanded.

The implementation can allow for different space dimensions for A and B, especially if the two coefficient matrices have rather different spectral properties. The idea of generating different approximation spaces—of the same dimension—for A and B by means of standard Krylov subspaces was first developed in [132], where, however, the right-hand side C of the original problem was approximated by a rank-one matrix $c_1 c_2^*$, with the aim of building the standard Krylov subspaces $K_j(A, c_1)$ and $K_j(B^*, c_2)$ as approximation spaces. The approach was then generalized to block Krylov subspaces in [225], so as to exploit the low- (but possibly larger than one) rank matrices C_1, C_2. Distinct Krylov subspaces for the right and left subspaces should also be considered when $B = A^*$, as long as $C_1 C_2^*$ is nonsymmetric. Nonetheless, in this case the generation of the two spaces can share some computationally intensive work, such as shifted system solves with the same coefficient matrix. The possibility of using nonsymmetric Lanczos processes which simultaneously generate $K_j(A, C_1)$ and $K_j(A, C_2)$ was explored in [135].

In Figure 1 we report a typical convergence history for the norm of the residual matrix, when the standard Krylov and extended Krylov subspaces are used for both A and B. Here A is the finite difference discretization of the Laplace operator in the unit square with homogeneous boundary conditions, and B is the same type of discretization for the operator $\mathcal{L}u = (\exp(-4xy)u_x)_x + (\exp(4xy)u_y)_y$, leading to matrices of the same size, $40{,}000 \times 40{,}000$; $C = c_1 c_2^*$, where c_1, c_2 are vectors with all entries equal to one, normalized to have unit norm. We note that with a subspace of dimension less than 120 for each matrix, the extended Krylov subspace is able to reduce the norm of the residual matrix by more than eight orders of magnitude, whereas the standard Krylov subspace of the same dimension shows almost no residual norm reduction. In terms of computational costs of the extended procedure, the matrices A and B are pretty sparse and systems involving them can be efficiently solved by a sparse direct solver.

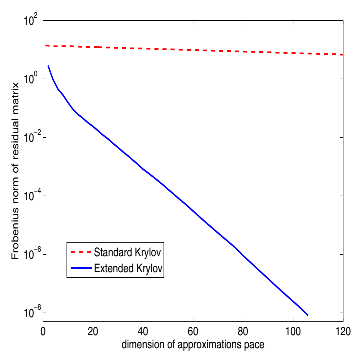

Fig. 1 *Typical convergence history of projection methods for the Sylvester equation with standard and extended Krylov subspaces.*

Rational Krylov subspaces have a rich history. First introduced by Ruhe in the context of eigenvalue approximation [212], their relevance has spread significantly in applied approximation theory and model order reduction frameworks due to their functional approximation properties; see, e.g., [4], [108], [114] and references therein.

The effectiveness of general rational spaces strongly relies on the efficiency of solving systems with A or its shifted variants. The reliability of recent direct sparse and iterative linear system solvers has made it possible to use these richer approximation spaces for more complex problems like the ones we are addressing. The choice of the shift is crucial to achieving fast convergence; this issue is postponed to the corresponding discussion for the Lyapunov equation in section 5.2.1.

In the quest for memory savings, the possibility of restarting the process could be considered: a maximum subspace dimension is allowed and the final approximate solution is obtained as $\widetilde{\mathbf{X}} = \widetilde{\mathbf{X}}^{(0)} + \widetilde{\mathbf{X}}^{(1)} + \widetilde{\mathbf{X}}^{(2)} + \cdots$, where the superscripts indicate a new restart. Strategies on how to generate the new approximations were proposed in [132]. We mention that new restarting procedures were recently proposed in [3], but their overall computational costs for large scale matrices have not clearly been assessed. An alternative that could be considered in the symmetric case is to resort to a two-pass strategy, inspired by a similar procedure in the eigenvalue context. Indeed, for A and B symmetric and not necessarily equal, an orthogonal basis of each standard Krylov subspace together with the projected matrix could be generated without storing the whole basis, but instead only the last three (block) vectors, because the orthogonalization process reduces to the short-term Lanczos recurrence [214]. Therefore, in a first pass only the projected solution $\mathbf{Y}$ could be determined while limiting the storage for V_k and W_j; at convergence the factors of the approximate solution $\widetilde{\mathbf{X}} = V_k \mathbf{Y} W_j^*$ could be recovered by generating the two bases once again. An implementation of such an approach can be found in [159] for $B = A^*$ and $C_1 = C_2$. The same idea could be used for other situations where a short-term recurrence is viable; the effectiveness of the overall method strongly depends on the affordability of computing the two bases twice.

The convergence analysis of projection methods has long been overlooked. Following recent significant advances in the convergence study of projection methods for the Lyapunov equation (see section 5.2.1), Beckermann in [20] provided a thorough study: residual norm bounds are given for Galerkin projection methods when rational Krylov subspaces, of possibly different dimensions, are used for A and B^*. The proposed estimates rely on new residual matrix relations and highlight the role of the field of values of the two coefficient matrices; we refer the reader to Proposition 5 below for further details on the results in [20]. Advances in the theoretical aspects of projection methods have been made in close connection with the recent progress in the understanding of polynomial and rational approximation methods for matrix functions such as the matrix exponential. The interplay of numerical linear algebra, approximation theory, and functional analysis has made this possible; see, e.g., [125], [114] and their references.

4.4.2. ADI Iteration. The ADI iteration was first introduced in [194] in 1955 and was proposed to solve large Sylvester equations by Ellner and Wachspress in [80].[11] Since then, and with various computationally effective refinements, the approach has been one of the leading methods for solving large scale Sylvester (and Lyapunov) equations. In its original form discussed in [80] and summarized next, the ADI iteration is derived for a full matrix $\mathbf{X}$ (see also Smith [233] for the derivation below). A low memory factorized version is used in practice for large matrices, and will be presented in what follows. In the following we assume that both real matrices A and B have eigenvalues with positive real parts. We can equivalently rewrite (18) as

$$(qI + A)\mathbf{X}(qI + B) - (qI - A)\mathbf{X}(qI - B) = 2qC, \quad q \neq 0.$$

For $q > 0$, $qI + A$ and $qI + B$ are nonsingular and we can multiply by their inverses to obtain the following equation:

$$\mathbf{X} - (qI + A)^{-1}(qI - A)\mathbf{X}(qI - B)(qI + B)^{-1} = 2q(qI + A)^{-1}C(qI + B)^{-1}.$$

Let $\mathcal{A} = (qI+A)^{-1}(qI-A)$, $\mathcal{B} = (qI-B)(qI+B)^{-1}$, and $\mathcal{C} = 2q(qI+A)^{-1}C(qI+B)^{-1}$. With this notation, the matrix equation above has the form $\mathbf{X} - \mathcal{A}\mathbf{X}\mathcal{B} = \mathcal{C}$ and is

[11] The authors of [80] referred to these Sylvester equations as Lyapunov equations.

called the Stein equation; see section 6. The matrix $\mathbf{X} = \sum_{k=1}^{\infty} \mathcal{A}^{k-1}\mathcal{C}\mathcal{B}^{k-1}$ is a formal solution to the Stein equation, and since both $\mathcal{A}$ and $\mathcal{B}$ have spectral radius less than one,[12] the series is convergent. This consideration drives the implementation of the following sequence of approximations:

$$\mathbf{X}_0 = \mathcal{C}, \quad \mathbf{X}_{k+1} = \mathcal{C} + \mathcal{A}\mathbf{X}_k\mathcal{B}. \tag{43}$$

The approach can be generalized to two parameters $p, q > 0$ for A and B, respectively, giving the transformed equation

$$\mathbf{X} - \mathcal{A}(p,q)\mathbf{X}\mathcal{B}(p,q) = \mathcal{C}(p,q),$$

with $\mathcal{A}(p,q) = (pI + A)^{-1}(A - qI)$, $\mathcal{B}(p,q) = (B - pI)(qI + B)^{-1}$, and $\mathcal{C}(p,q) = (p+q)(pI + A)^{-1}C(qI + B)^{-1}$. A recursion similar to the one for a single parameter can be derived, and it is convergent if the spectral radii of $\mathcal{A}(p,q), \mathcal{B}(p,q)$ are both less than one. Therefore, the parameters p, q are selected so as to minimize these spectral radii, and if A, B are both symmetric with spectral intervals $[a, b]$ and $[c, d]$, respectively, this corresponds to solving the ADI minimax problem

$$\min_{p,q>0} \max_{s\in[a,b],t\in[c,d]} \left| \frac{(q-s)(p-t)}{(p+s)(q+t)} \right|.$$

The generalization of this concept allows one to choose different p, q at each iteration, allowing for a sequence of parameters $p_1, p_2, \ldots$ and $q_1, q_2, \ldots$. The associated ADI minimax problem after J iterations thus becomes

$$\min_{p_j,q_j>0} \max_{s\in[a,b],t\in[c,d]} \prod_{j=1}^{J} \left| \frac{(q_j-s)(p_j-t)}{(p_j+s)(q_j+t)} \right|,$$

which, if solved exactly, provides optimal parameters for the convergence rate of the ADI iteration; that is, theoretical convergence rates can be achieved for matrices with real spectra [255]. In practice, a fixed number J of parameters is selected a priori and then cyclically repeated until convergence. The choice of J is driven by the quality of the computed parameters: fewer parameters may be better than many badly distributed parameters. We will return to this issue in section 5.2.

Following a successful idea developed for the Lyapunov equation, the authors of [31] proposed a factorized version of the ADI iteration, which allows one to write the approximate solution as the product of three memory saving factors, as long as $C = C_1C_2^*$ is low rank. We will expand on this and other implementation aspects such as rank truncation in the case of the Lyapunov equation, since the changes occurring when generalizing ADI to the Sylvester equation are mainly technical and due to the presence of two distinct approximation spaces; we point here to the recent work of Peter Benner and his collaborators for a comprehensive implementation investigation of ADI for the Sylvester equation.

We conclude with a theoretical comparison recently made between ADI and the Galerkin method (see section 4.4.1) in the rational Krylov subspaces

$$K_m(A, C_1) = \text{range}\{C_1, (A + \sigma_2 I)^{-1}C_1, \ldots, (A + \sigma_m I)^{-1}C_1\},$$

$$K_m(B^*, C_2) = \text{range}\{C_2, (B^* + \eta_2 I)^{-1}C_2, \ldots, (B^* + \eta_m I)^{-1}C_2\},$$

[12] For a given matrix A with eigenvalues λ in $\mathbb{C}^+$ and $q > 0$, the eigenvalues of $(qI - A)(qI + A)^{-1}$ are given by $(q - \lambda)/(q + \lambda)$, with absolute values all less than one.

where in both cases the first pole is taken to be at infinity, so that the columns of C_1 and C_2 belong to the corresponding spaces. In [88], Flagg and Gugercin showed that ADI and the Galerkin approach are equivalent whenever the poles of both methods coincide with the eigenvalues of the projections of A and B (Ritz values) in the two spaces, respectively; the same result was earlier proved for the Lyapunov equation with different techniques (see Theorem 11). Moreover, for general poles the following result was proved by Beckermann for the error [20, Cor. 2.2].

PROPOSITION 5. *Let* $\mathbf{X}$ *be the exact solution to the Sylvester equation. Let* $\mathcal{S}_{A,B}(X) = AX + XB$, *and let* $\mathbf{X}^G_{k,k}$, $\mathbf{X}^{ADI}_k$ *be the approximate solutions obtained after k iterations of the Galerkin method in* $K_k(A, C_1), K_k(B^*, C_2)$ *and after k ADI steps, respectively, with the two methods using the same poles. If the fields of values* $W(A)$ *and* $W(-B)$ *have empty intersection, then*

$$\|\mathcal{S}_{A,B}(X - X^G_{k,k})\|_F \le \gamma_0 \|\mathcal{S}_{A,B}(X - X^{ADI}_k)\|_F$$

with constant $\gamma_0 \le 3 + 2c_0$, *with* $c_0 = 2\mathrm{diam}(W(A), W(-B))/\mathrm{dist}(W(A), W(-B))$ *independent of the poles used to generate the space.*

The constant γ_0 is not optimal. As stated in [20], Proposition 5 shows that, even for optimal poles, ADI cannot give much better results than rational Galerkin; moreover, for poor poles ADI is known to give much larger residuals. Further results will be discussed for the case of the Lyapunov equation.

4.4.3. Data Sparse and Other Methods. A variety of approaches relying on the data sparsity structure has been analyzed. These methods may be particularly appropriate in the large scale case when the right-hand side matrix C is sparse and full rank.

The Kronecker formulation allows one to consider a wide range of linear system solvers for (20); an early ad hoc implementation of the classical SOR was proposed in [237], although the exploding dimensions of the problem significantly penalize the method when compared with the approaches analyzed so far. We also recall from section 4.4.1 that global Krylov subspace methods represent an implicit way to deal with the Kronecker formulation. Other iterative solvers based on the Kronecker formulation (20) have been explored specifically for the Lyapunov equation, and they will be reviewed in section 5.2.3. These appear to be the main directions taken whenever C is not numerically low rank.

For data sparse matrices $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$, namely, such that matrix-vector multiplications for A and B can be performed with complexity $\mathcal{O}(n)$ and $\mathcal{O}(m)$, respectively, an extension of the multigrid algorithm was more recently proposed in [104], where A and B stem from the discretization of a class of partial differential equations and their spectra are assumed to be separated by a line. A particular computational caveat of this extension is the smoother (e.g., Jacobi), which in this case requires approximately solving a *diagonal* Sylvester equation at each iteration. This step is carefully discussed in [104], and a procedure for determining a cheap and low-rank approximate solution is devised. Other crucial aspects include handling the connection between the (independently generated) sequences of matrices for A and B, which is accounted for during the smoothing procedure, and the imposition of regularity constraints on the continuous operators associated with A and B. A major issue arising when using these hierarchical methods is whether the approximate solution $\widetilde{\mathbf{X}}$ is low rank, so that it can be stored cheaply by means of a (hierarchical) sparse format, the $\mathcal{H}$-matrix format. Such a format is a data sparse representation for a special class of matrices which appear to occur after the use of several discretization methods, when partical differential equations or integral equations are treated numerically [101]. The $\mathcal{H}$-matrix format consists of partitioning a given matrix recursively into submatrices admitting low-rank approximations. The definition of this format requires the introduction of further arithmetic operations/approximations, in order to be able to determine, e.g., an $\mathcal{H}$-matrix after the approximate inversion of an $\mathcal{H}$-matrix, in order to make the class closed with respect to some important matrix operations; see section 5.2.3 for further details.

A different though related approach consists in adapting small scale iterations to the large setting, again under the condition that C is low rank. This can be performed, for instance, within the sign function iteration, by using rank truncation of the iterates and sparse format for the approximate solution. More details on the sign function iteration will be given in section 5.2.3. Here we mention that such an approach is investigated in [17] (see also [19]), where the sparse format chosen for the data and for the approximate solution is the hierarchical $\mathcal{H}$-matrix format also used in [102], [104]. With this approach, sparse approximate solutions to a Sylvester equation of size up to $n = 262144$ associated with a control problem for the two-dimensional heat equation are reported in [17]. The accuracy and effectiveness of the method depend on some thresholds used for maintaining sparsity and low rank during the iteration, and are thus problem dependent.

5. Continuous-Time Lyapunov Equation.

For $B = A^*$, from the Sylvester equation we obtain the Lyapunov equation

$$A\mathbf{X} + \mathbf{X}A^* + C = 0, \tag{44}$$

with C symmetric, and its generalized counterpart $A\mathbf{X}E^* + E\mathbf{X}A^* + C = 0$, with E nonsingular. Clearly, this latter equation can be transformed into the form (44) by left and right multiplication by E^{-1} and E^{-*}, respectively. If E is symmetric and positive definite, a Cholesky decomposition could be performed and its inverse factors applied to the equation on the left and right sides, to maintain the problem structure. These are called the *continuous-time* Lyapunov equations, to be distinguished from the discrete-time equations which will be discussed in section 6. They arise in the analysis of continuous-time and discrete-time linear dynamical systems, respectively. A very detailed analysis of the Lyapunov equation, with computational developments up to 1995 and many relevant connections in the control application area, can be found in [90].

In the context of inertia theory, (44) with $C \succeq 0$ relates the location of the eigenvalues of A and $\mathbf{X}$ with respect to the imaginary axis. Since C is symmetric, the solution $\mathbf{X}$ is also symmetric. According to the Sylvester equation theory, the solution to (44) exists and is unique if and only if $\lambda_i + \bar{\lambda}_j \neq 0$ for all eigenvalues λ_i, λ_j of A [131]. If all eigenvalues of A have negative real part, namely, A is stable, then this condition is satisfied, so that a unique solution is ensured. We remark that the stability of A is an important property in the control setting, therefore it is not regarded as a restriction for solving the Lyapunov equation, although not strictly required. We shall see, however, that some of the large scale methods require additional restrictions on A, namely, its negative definiteness, to ensure the existence of an approximate solution. For A nonsymmetric, this extra condition may limit the applicability of the method, since in general a stable matrix A is not necessarily negative definite.

It can be verified that if A is stable and $C \succ 0$ ($C \succeq 0$), then $\mathbf{X} \succ 0$ ($\mathbf{X} \succeq 0$); in this case the problem is called the *stable* Lyapunov equation. If $C \succeq 0$ and (A, C^*) is observable, then $\mathbf{X}\succ 0$. A detailed account of various relations between the inertia of A and that of $\mathbf{X}$ can be found, e.g., in [168, sec. 13.1], [221], [222]. A specialized sensitivity bound can be obtained for the stable Lyapunov equation. Assume that $\mathbf{X} + \Delta\mathbf{X}$ solves

$$(A + \Delta A)(\mathbf{X} + \Delta\mathbf{X}) + (\mathbf{X} + \Delta\mathbf{X})(A + \Delta A)^* + (C + \Delta C) = 0;$$

then

$$\frac{\|\Delta\mathbf{X}\|}{\|\mathbf{X} + \Delta\mathbf{X}\|} \leq 2\|A + \Delta A\| \, \|H\| \left[\frac{\|\Delta A\|}{\|A + \Delta A\|} + \frac{\|\Delta C\|}{\|C + \Delta C\|}\right],$$

where H satisfies $AH + HA^* + I = 0$ and all denominators are assumed to be nonzero [121]. Estimates for the backward error associated with the Lyapunov equation do not differ from those in (27) for the Sylvester equation; therefore, except for the substitution $B = A^*$, the extra structure of the problem does not modify the sensitivity properties of the solution [123].

The sensitivity of the solution to (44) can also be analyzed by looking at the spectral properties of the solution matrix; this topic has attracted a lot of interest, especially in light of its consequences for the stability analysis of dynamical systems. Various authors have explored the spectral decomposition of the Lyapunov solution to analyze the decay of its eigenvalues; see, e.g., [199], [235], [5], [155]. In [5], an error estimate for a low-rank approximation to the solution of (44) was proved. For the sake of simplicity we report here only the case when C is rank one. The result relies on the fact that the solution matrix admits the following decomposition:

$$\mathbf{X} = ZDZ^*, \quad D = \mathrm{diag}(\delta_1, \ldots, \delta_n), \quad \delta_k = \frac{-1}{2\Re(\lambda_k)} \prod_{j=1}^{k-1} \left|\frac{\lambda_k - \lambda_j}{\bar{\lambda}_k + \lambda_j}\right|^2,$$

where λ_j are the eigenvalues of the diagonalizable matrix A.

THEOREM 6. *Assume A is diagonalizable with eigenvector matrix Q having all unit norm columns, and let $C = cc^*$. Let $\mathbf{X} = \sum_{j=1}^n \delta_j z_j z_j^*$ solve* (44)*, with the nonnegative values δ_j sorted decreasingly, and for $k \in \{1, \ldots, n\}$ define $\mathbf{X}_k = \sum_{j=1}^k \delta_j z_j z_j^*$. Then*

$$\|\mathbf{X} - \mathbf{X}_k\| \leq (n-k)^2 \delta_{k+1} (\kappa(Q)\|c\|_2)^2,$$

where $\| \ \|$ is the matrix norm induced by the vector 2-norm.

The bound may not be sharp for highly nonnormal A, for which $\kappa(Q)$ may be large. A more specialized bound was earlier given by Penzl for A symmetric, which depends only on the condition number of A [199].

THEOREM 7 (see [199]). *Let A be symmetric and negative definite, with condition number $\kappa(A)$, and let $C = C_1 C_1^*$ with C_1 of rank p. Let $\lambda_i(\mathbf{X})$ with $i = 1, \ldots, n$ be the nonincreasingly ordered eigenvalues of $\mathbf{X}$. Then*

$$\frac{\lambda_{pk+1}(\mathbf{X})}{\lambda_1(\mathbf{X})} \leq \left(\prod_{j=0}^{k-1} \frac{\kappa^{(2j+1)/(2k)} - 1}{\kappa^{(2j+1)/(2k)} + 1}\right)^2 \quad \text{for} \quad 1 \leq pk < n.$$

It was also shown in [199] that eigenvalues alone cannot predict the eigenvalue distribution of $\mathbf{X}$ in the non-Hermitian case; this is reminiscent of similar limitations of nonnormal matrices in the convergence analysis of iterative methods for linear systems [176, sec. 5.7.3].

Bounds on the eigenvalue decay that attempt to cope with nonnormality were obtained in [215, sec. 3.1.2], where the concept of pseudospectrum is used; there, some interesting counterintuitive convergence behaviors are also described. Overall, much remains to be understood about the decay of the solution spectrum in the nonnormal case.

In addition to the application relevance, establishing conditions under which the solution matrix has exponentially decaying eigenvalues provides theoretical grounds for the good performance of low-rank methods in the large scale case.

5.1. Lyapunov Equation. Small Scale Computation. As for the Sylvester equation, the closed form solutions described in section 4 could be used in theory. A detailed account of early methods can be found in [90], together with some ad hoc algorithms appropriate when special forms of A—e.g., Schwarz, companion, or Jordan forms—are available; see also [38], [120] for improved approaches for the companion form.

The standard method for efficiently solving (44) when A has small dimensions does not essentially differ from those for the Sylvester equation discussed in previous sections. In fact, since $B = A^*$, the computational cost of the reduction to Schur form is halved in the Bartels–Stewart method [223].

A specifically designed algorithm was proposed by Hammarling to exploit the case when C is positive semidefinite. It was shown in [116] that if $C = C_1 C_1^* \succeq 0$, it is possible to determine the Cholesky factor L of the solution $\mathbf{X} = LL^*$ without first determining $\mathbf{X}$. The computation of the Cholesky factor has some advantages when $\mathbf{X}$ is nonsingular but severely ill-conditioned, as it is the case when the singular values decay rapidly, so that dealing with L significantly improves the accuracy and robustness of computations with $\mathbf{X}$; in [272] a comparison between Hammarling's method and the Bartels–Stewart method can be found. A block variant of Hammarling's method for the discrete-time Lyapunov equation is suggested in [158], which can dramatically improve the performance of the original scalar (unpartitioned) algorithm on specific machine architectures, while preserving the stability of the original method.

We also mention the possibility of preprocessing, in both the continuous- and the discrete-time equations, in order to transform the original symmetric problem into a skew-symmetric one, so that the solution will also be skew-symmetric ($\mathbf{X} = -\mathbf{X}^*$), allowing for some memory savings; see [90, sec. 2.1.2] and references therein.

A completely different approach exploits the fact that the solution $\mathbf{X}$ may be computed by means of matrix functions, in particular, by using the sign function. Although less general than Schur-form-based algorithms, they allow one to handle larger problems, especially if the right-hand side is low rank or structured, and can be more easily adapted to a high performance computational environment. The idea is to use well-established matrix iterations to obtain the matrix sign function in a cheap manner by fully exploiting the possible sparse format of the matrix. The whole procedure is actually more general, and it also applies to the symmetric algebraic Riccati equation (see Algorithm 1). Here we will follow the derivation proposed in [18] (see also [33]), although the main iteration was introduced by Larin and Aliev in [169] for the generalized Lyapunov equation. Let $\mathcal{A} = \mathcal{X}\text{blkdiag}(J_+, J_-)\mathcal{X}^{-1}$ be the Jordan decomposition of a given matrix $\mathcal{A}$, where J_+, J_- represent the Jordan matrices associated with the eigenvalues in the open planes $\mathbb{C}^+$ and $\mathbb{C}^-$, respectively; the decomposition thus assumes that $\mathcal{A}$ has no purely imaginary eigenvalues. Then $\text{sign}(\mathcal{A}) = \mathcal{X}\text{blkdiag}(I, -I)\mathcal{X}^{-1}$, where the dimensions of I and $-I$ match those of J_+ and J_-, respectively. For A stable, the solution to the Lyapunov equation satisfies (see, e.g., [210])

$$\begin{bmatrix} 0 & \mathbf{X} \\ 0 & I \end{bmatrix} = \frac{1}{2}\left(I + \text{sign}\left(\begin{bmatrix} A^* & C \\ 0 & -A \end{bmatrix}\right)\right) =: \frac{1}{2}\left(I + \text{sign}\left(Z_0\right)\right). \tag{45}$$

With this property, the following matrix iteration corresponds to applying the Newton method to the nonlinear equation $(\text{sign}Z_0)^2 = I$:

$$Z_{k+1} = \frac{1}{2}(Z_k + Z_k^{-1}), \qquad k = 0, 1, \ldots. \tag{46}$$

This yields

$$\text{sign}\, Z_0 = \lim_{k\to\infty} Z_k = \begin{bmatrix} -I & 2\mathbf{X} \\ 0 & I \end{bmatrix}.$$

Although the iteration is globally and (asymptotically) quadratically convergent, the basic iteration above may have slow initial convergence, so it is often accelerated using a parameterized procedure, that is, $Z_{k+1} = \frac{1}{2}(c_k Z_k + (c_k Z_k)^{-1})$, $k = 0, 1, \ldots$, for an appropriate selection of the parameter $c_k > 0$. A popular choice is $c_k = |\det(Z_k)|^{-\frac{1}{n}}$ [50]; see, e.g., [7], [42, sec. 3.5.2] for a review of other choices and for additional historical and computational considerations on the matrix sign function.

5.2. Lyapunov Equation. Large Scale Computation. Recalling the discussion for the Sylvester equation in section 4.4, the solution of the Lyapunov equation for A of large dimensions focuses on the determination of memory saving and computationally appealing approximations. For the stable problem, this is achieved in most cases by looking for a *low*-rank approximation $\widetilde{\mathbf{X}} = \mathbf{Z}\mathbf{Z}^*$, so that only the tall matrix $\mathbf{Z}$ is actually computed and stored. This can be possible if, for instance, the right-hand side has low rank, since in that case we also have $\mathbf{X} \succeq 0$. Nonetheless, strategies to approximate the general right-hand side by low-rank matrices have also been explored in the literature; see, e.g., [132].

To help fully grasp the relevance of the topic, we notice that a number of recent Ph.D. theses have been devoted to the theory and computational aspects of the large scale Lyapunov matrix equation, and their results have significantly advanced knowledge on the problem; among them, we note [197], [189], [185], [129], [272], [173], [215]. The list could be expanded if one were to also include closely related theses on model order reduction of linear dynamical systems.

We conclude this section by noting that a systematic numerical comparison of all iterative methods described in the following subsections on a variety of very large problems (of size $n \gg 10^4$) is still lacking, although in our presentation some guidelines are given about the settings in which each of the methods discussed is preferred.

5.2.1. Projection Methods. As in the case of the Sylvester equation, the derivation of a projection method can be determined by imposing, e.g., the Galerkin condition on the residual with respect to some approximation space. In particular, from (41) with $k = j$, $V_k = W_j$, and $C_2 = C_1$, we obtain the projected small size Lyapunov equation

$$V_k^* A V_k \mathbf{Y}_k + \mathbf{Y}_k V_k^* A^* V_k + V_k^* C_1 (V_k^* C_1)^* = 0, \tag{47}$$

whose solution matrix $\mathbf{Y}_k$ gives $\mathbf{X}_k = V_k \mathbf{Y}_k V_k^* \approx \mathbf{X}$. Since $\mathbf{Y}_k$ is positive semidefinite and numerically singular, it is possible to perform a truncated decomposition of $\mathbf{Y}_k$ as $\mathbf{Y}_k = LL^*$, so that only the slim factor $Z_k = V_k L$ of the solution $\mathbf{X}_k = Z_k Z_k^*$

needs to be stored. To ensure that (47) admits a unique solution, the matrix $V_k^* A V_k$ is assumed to be stable. Such a sufficient condition is met by requiring that A be negative definite, which is the usual hypothesis when using projection methods. This condition represents a limitation of projection methods, since the original problem admits a unique solution even in case of a stable[13] but not necessarily negative definite A. On the other hand, these are sufficient conditions: projection methods can work in practice without this assumption, although they may break down or show some erratic convergence behavior; see [179] for an analysis.

An apparently different (functional) approach, based on the approximation to the matrix exponential and on (21), leads to exactly the same approximation procedure as Galerkin. Indeed, the action of the matrix exponential on C_1, $\exp(tA)C_1$, can be approximated in the space $\mathcal{V}$ as $V_k \exp(tH_k)(V_k^* C_1)$, where $H_k = V_k^* A V_k$, so that the analytic expression in (21) for the solution can be approximated explicitly; this is the way the Galerkin approximate solution was originally obtained in [213] for a rank-one matrix C_1.

PROPOSITION 8 (see [213]). *Let $\mathcal{V}$ be a subspace of $\mathbb{R}^n$, and let V have orthonormal columns and be such that $\mathcal{V} = \mathrm{range}(V)$. Let H be the projection and restriction of A onto $\mathcal{V}$, and $y(t) = \exp(tH)(V^* C_1)$. Then the matrix $V\mathbf{Y}V^*$ with*

$$\mathbf{Y} = \int_0^\infty y(t)y(t)^* dt$$

is the Galerkin approximate solution to the Lyapunov equation in $\mathcal{V}$.

The procedure above is very general, and the success of the approach, in terms of computational cost, depends on the choice of the approximation space $\mathcal{V}$. All choices discussed in section 4.3 have been explored. For instance, the block Krylov subspace $K_k^\square(A, C_1)$ was exploited in [137] and was referred to as the Arnoldi method, after the procedure used to build the block Krylov subspace. The following computationally convenient relation for the residual $R_k = A\mathbf{X}_k + \mathbf{X}_k A^* + C_1 C_1^*$ can be deduced from (42) [137, Thm. 2.1]:

$$\|R_k\|_F = \sqrt{2}\|(v_{k+1}^* A v_k) E_k^* \mathbf{Y}_k\|_F, \qquad E_k^* = [0_m, \ldots, 0_m, I_m],$$

where v_{k+1} contains the next block of basis vectors.

Finally, the solution $\mathbf{X}_k$ is the exact solution to the nearby problem [137]

$$(A - \Delta)\mathbf{X} + \mathbf{X}(A - \Delta)^* + C_1 C_1^* = 0,$$

with $\Delta = V_{k+1}(V_{k+1}^* A V_k)V_k^*$, $\|\Delta\|_F = \|V_{k+1}^* A V_k\|_F$.

The asymptotic convergence of the Arnoldi method was recently analyzed in [228]. Here we give an example of such analysis, which applies to A symmetric and positive definite and C_1 of rank one and unit norm; the derived bound was shown in [228] to provide an accurate worst-case convergence rate of the method.

THEOREM 9. *Let A be symmetric and positive definite, and let $\lambda_{\min}$ be the smallest eigenvalue of A. Let $\hat\lambda_{\min}$, $\hat\lambda_{\max}$ be the extreme eigenvalues of $A + \lambda_{\min} I$ and $\hat\kappa = \hat\lambda_{\max}/\hat\lambda_{\min}$. Let $\mathbf{X}_k$ be the Galerkin approximate solution to $\mathbf{X}$ in a Krylov*

[13] As said before, even stability of A is not strictly necessary for the solvability of the Lyapunov equation, only that $I \otimes A + A \otimes I$ is nonsingular.

subspace of dimension k. Then

$$\|\mathbf{X} - \mathbf{X}_k\| \le \frac{\sqrt{\hat{\kappa}}+1}{\hat{\lambda}_{\min}\sqrt{\hat{\kappa}}} \left(\frac{\sqrt{\hat{\kappa}}-1}{\sqrt{\hat{\kappa}}+1} \right)^k, \tag{48}$$

where the matrix norm is that induced by the vector 2-norm.

This result shows that the error norm is bounded by the same asymptotic quantity as for the conjugate gradient method applied to a standard linear system with coefficient matrix $A + \lambda_{\min} I$.

As mentioned above, the algorithmic steps to computing an approximate solution by projection remain unchanged when a different approximation space is used. In [226] an efficient method based on the extended Krylov subspace in (38), namely, $K_k^{\square}(A, C_1) + K_k^{\square}(A^{-1}, A^{-1}C_1)$, was introduced for C_1 of low rank. In [226] the method was shown experimentally to be highly superior to the Arnoldi method for sparse and large coefficient matrices, allowing the computation of an equally accurate solution with a significantly smaller dimensional subspace, at lower computation costs for sparse A; convergence plots are typically similar to that in Figure 1. According to the experiments in [226], the method also compares well with respect to ADI. A recent asymptotic analysis in [153] theoretically confirmed these results, showing that the extended Krylov subspace method is expected to have higher convergence rate than the Arnoldi method and than ADI with a single pole.

More recently, rational Krylov subspaces have been introduced as approximation spaces in the solution of the Lyapunov equation, showing the great potential of projection methods when solving shifted systems with A is affordable [77]. In [76] estimates for the error norm were derived, assuming that the field of values $W(A)$ of A and the set of parameters $\{\sigma_i\}_{i=1,2,\ldots}$ are both contained in the same half-plane. We first report on an error bound that emphasizes the connection with ADI.

THEOREM 10. *Let $C_1 \in \mathbb{R}^n$, and let $\mathbf{X}_k$ be the approximation obtained by a Galerkin method in range $([(A+\sigma_1 I)^{-1}C_1, \ldots, (A+\sigma_k I)^{-1}C_1])$. Let $\gamma_1 > 0$ be the Crouzeix constant (with $\gamma_1 \le 11.08$) and $\gamma_2 = \frac{1}{2\mathrm{dist}(0, W(A))}$. Then*

$$\|\mathbf{X} - \mathbf{X}_k\|_F \le 2{\gamma_1}^2 \gamma_2 \max_{z \in W(A)} \frac{\prod_i |z - \bar{\sigma}_i|^2}{\prod_i |z + \sigma_i|^2} \|C_1\|^2, \tag{49}$$

where $\|\cdot\|_F$ is the Frobenius norm.

The maximization problem appearing in the upper bound is the same as that characterizing the convergence rate in the ADI method (see section 5.2.2). We also note that the ADI bound is reachable, in terms of convergence rate, therefore it may be viewed as a worst-case scenario for ADI. Therefore, the result of Theorem 10 provides a realistic picture of the performance of the rational Krylov subspace method compared with that of ADI, whose implementation for the Lyapunov equation is described in detail in section 5.2.2. The superiority of the rational Krylov subspace method for the same shifts can be easily appreciated in practical cases; we refer the reader to, for instance, Examples 4.3–4.4 and Figure 4.5 in [76]. Beckermann in [20] expanded this type of result to the more general setting of the Sylvester problem (see Proposition 5). Nonetheless, if close to optimal poles can be selected, ADI may still provide a competitive alternative to rational Krylov subspace methods.

Asymptotic error bounds for the Galerkin method in the rational Krylov subspace were derived in [76]. The reported numerical experiments on worst-case spectral distributions show that these bounds are indeed sharp for certain classes of data.

The general rational Krylov subspace requires the selection of a series of shifts (poles), which can either be computed a priori or during the generation of the space. The a priori pole computation may possibly incur high computational costs, following the same procedure as that used for other parameter-dependent methods such as ADI (see below). On the other hand, it was shown in [77] that it is possible to employ a greedy algorithm to compute the next pole on the fly, while the iteration proceeds, with computational costs of a modest power of the order of the space dimension, which is usually significantly smaller than the problem dimension. This is done by exploiting approximate spectral information generated within the current approximation space. More precisely, for $C_1 = c_1 \in \mathbb{R}^n$, we first observe that an element of the rational Krylov subspace of dimension k can be written as $x = p_{k-1}(A)q_k(A)^{-1}c_1$ with p_{k-1} and q_k polynomials of degree $k-1$ and k, respectively, where the roots of q_k are the parameters $\sigma_1, \ldots, \sigma_k$. For the sake of the derivation, assume that the linear system $(A + sI)x = c_1$ for some parameter $s > 0$ is to be solved. Then the residual of an approximate solution x_k obtained by imposing the Galerkin condition with respect to the space can be written as

$$c_1 - (A + sI)x_k = \frac{r_k(A)c_1}{r_k(s)}, \qquad r_k(z) = \prod_{j=1}^{k} \frac{z+\lambda_j}{z+\sigma_j},$$

where $\lambda_1, \ldots, \lambda_j$ are the eigenvalues of the projection of A onto the space (Ritz values). The adaptive procedure amounts to determining the next parameter σ_{k+1} by capturing the parameter σ for which the current rational function is largest:

$$\sigma_{k+1} = \arg\left(\max_{\sigma\in\partial S_k} \frac{1}{|r_k(\sigma)|}\right),$$

where $S_k \subset \mathbb{C}^+$ approximates the spectral region of A and ∂S_k is its border. The actual computational procedure requires an initial rough estimate of ∂S_k, which for real A can be taken to be the approximate extreme values of the interval $S_k \cap \mathbb{R}$. These can be easily obtained by a few iterations of an eigenvalue solver [77].

Numerical experiments reported in [77] show that the method is superior to the extended Krylov subspace when, for instance, the field of values of A is very close to the imaginary axis. The computational cost of the general rational Krylov subspace method may be much higher than for the extended space, since a group of new shifted linear systems with the same right-hand side needs to be solved at each iteration, with a different shift at each one. On the other hand, the extended method only requires system solves with A: if A is such that an efficient LU decomposition can be performed, then this is done once at the beginning of the computation and only the backward solves are required while expanding the space. The numerical experiments reported in [77] suggest that the rational approximation space dimension usually remains very low, so that only a few systems have to be solved. The difference in computational costs per iteration is less significant if an iterative solver is used to solve the inner systems; in that case, the extended method can possibly still reuse the same preconditioner, but the computation with the iterative method still needs to be performed anew.

The rational function idea is particularly appealing when $C = C_1C_1^*$ has rank p larger than one. In that case, the extended Krylov subspace increases its dimension by $2p$ vectors per iteration, making the whole procedure memory consuming if convergence is not fast.

In general, memory requirements may become a serious concern when C_1 has rank much larger, say a few dozen, since the approximation space dimension increases proportionally with that rank. In [78] a tangential procedure is proposed to expand the rational Krylov subspace at each iteration in a way that ensures only the most relevant directions are retained.[14] More precisely, small tall matrices $d_1, \dots, d_k$ are determined so that the space is expanded only with the linear combinations $C_1 d_1, \dots, C_1 d_k$, giving

$$\mathrm{range}([(A+\sigma_1 I)^{-1}C_1 d_1, (A+\sigma_2 I)^{-1}C_1 d_2, \dots, (A+\sigma_k I)^{-1}C_1 d_k]).$$

The actual column dimension of each d_j, between 1 and p, may vary as the iteration proceeds. The next pair (d_{k+1}, σ_{k+1}) is obtained on the fly by an optimization procedure. Numerical experiments reported in [78] show that this strategy is capable of successfully handling the presence of many columns in C_1 and provides a performance that is largely superior to that of the block rational Krylov subspace.

The global Krylov subspace method for the Sylvester equation was applied to the Lyapunov equation in [142], with natural simplifications due to the fact that a single space needs to be generated; numerical experiments in [142] showed better performance than the standard block Krylov subspace methods. No numerical comparisons seem to be available for global and rational Krylov subspace methods. We also mention [122] for a generalization to the simultaneous solution of a coupled pair of Lyapunov equations, corresponding to the two Gramians of a dynamical system. This last problem was also considered in [135]: the coupled block Arnoldi method and the nonsymmetric block Lanczos recurrence were analyzed as candidates for simultaneously approximating both Gramians in order to obtain approximations to the linear transfer function of the system; see also [136] for enhancements of the proposed approaches.

The Galerkin condition for the Lyapunov equation residual can be replaced by a Petrov–Galerkin condition; see the discussion around Proposition 3 for the Sylvester equation. If the constraint space is $A\mathcal{V}$, then the resulting algorithm minimizes the residual in the Frobenius norm and the problem admits the following formulation: Find $\mathbf{X}_k = V_k \mathbf{Y}_k V_k^*$ such that $\mathbf{Y}_k$ satisfies

$$\mathbf{Y}_k = \arg\min_{Y_k \in \mathbb{R}^{pk\times pk}} \|AV_k Y_k V_k^* + V_k Y_k V_k^* A^* + C_1 C_1^*\|_F, \tag{50}$$

where the columns of V_k form a basis for the approximation space $\mathcal{V}$. Assume once again that for the space $\mathcal{V}$ a relation of the type $AV_k = \widehat{V}_{k+1}\underline{H}_k$ is available, with $\widehat{V}_{k+1}^*\widehat{V}_{k+1} = I$ and $\underline{H}_k$ of size $p(k+1)\times pk$. Then (50) can be rewritten in smaller dimension as the following matrix least squares problem (see, e.g., [179]):

$$\mathbf{Y}_k = \arg\min_{Y \in \mathbb{R}^{pk\times pk}} \left\| \underline{H}_k Y \begin{bmatrix} I, & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} Y \underline{H}_k^* + \begin{bmatrix} \gamma_0^2 & 0 \\ 0 & 0 \end{bmatrix} \right\|_F, \tag{51}$$

where $C_1 = V_k\gamma_0$. This approach was explored in [137] for the standard block Krylov subspace and in [132] for the rank-one case. The projected problem entails the solution of (51) for which expensive (order $\mathcal{O}((pk)^4)$) procedures have been proposed [137], [132]. More recently, the minimal residual method was revisited in [179] for a generic low-rank C_1, and a more effective (order $\mathcal{O}((pk)^3)$) solver for the inner problem (51) was proposed.

[14]The term "tangential" comes from first-order tangential interpolation properties of these spaces in the context of model order reduction [78].

5.2.2. ADI Method. For $B = A^*$, the ADI method of section 4.4.2 for the Sylvester equations is simplified, leading to the following recursion for the whole matrix $\mathbf{X}_j$ with two half steps (see [255]):

$$\begin{aligned} \mathbf{X}_0 &= 0, \\ (A + s_j I)\mathbf{X}_{j-\frac{1}{2}} &= -C_1 C_1^* - \mathbf{X}_{j-1}(A^* - s_j I), \\ (A + s_j I)\mathbf{X}_j &= -C_1 C_1^* - (\mathbf{X}_{j-\frac{1}{2}})^*(A^* - s_j I), \quad j = 1, \ldots, k. \end{aligned}$$

Here the shifts $\{s_j\}$ are complex and are employed cyclically. If both A and C_1 are real, then the approximate solution will be real and symmetric as long as both complex conjugates shifts are used [181]. A key idea to make the recursion amenable to large dimension matrices is to keep the solution iterate in factored form. This idea was successfully explored by Penzl in 2000 in [198] and was the basis for the software package `Lyapack` [200]; see also [30]. The resulting low-rank ADI (LR-ADI) method thus determines a recurrence for the factor $\mathbf{Z}_j$ of $\mathbf{X}_j = \mathbf{Z}_j \mathbf{Z}_j^*$ as

$$\mathbf{Z}_{j+1} = [(A^* - s_j I)(A^* + s_j I)^{-1}\mathbf{Z}_j, \sqrt{-2s_j}(A^* + s_j I)^{-1}C_1], \tag{52}$$

with $\mathbf{Z}_1 = \sqrt{-2s_1}(A^* + s_1 I)^{-1}C_1$: the number of columns in the factor $\mathbf{Z}_j$ is enlarged by $\operatorname{rank}(C_1)$ columns at each iteration. The success of LR-ADI is related to what Penzl called the *low-rank phenomenon* in the solution $\mathbf{X}$: namely, the previously mentioned fact that the eigenvalues of $\mathbf{X}$ tend to decay very quickly toward machine precision, so that a low-rank approximation is possible (see section 4.1).

The iteration matrix $\mathbf{Z}_j$ is complex during the whole iteration, whenever some of the shifts are complex. A way to overcome this problem and to maintain real arithmetic throughout whenever A and C_1 are real is discussed in [198]; see also the more recent contribution [28].

The iteration in (52) requires solving systems with right-hand sides $\mathbf{Z}_j$ *and* C_1 at each step j. A computational improvement to decrease the number of solves per iteration was suggested in [174] (where the LR-ADI method was called CF-ADI, in which CF stands for Cholesky Factor). There, the columns were reordered so that only the previous iterate requires solves with a shifted matrix. The resulting recurrence is given in the following algorithm (see [174, Alg. 2]):

Algorithm 6. Given A, C_1, and $\{s_j\}$, $j = 1, \ldots, j_{\max}$:

1. Set $\mathbf{z}_1 = \sqrt{-2s_1}(A + s_1 I)^{-1}C_1$, $\mathbf{Z}_1 = \mathbf{z}_1$.
2. For $j = 2, \ldots, j_{\max}$
   - 2.1. $\mathbf{z}_j = \frac{\sqrt{-2s_j}}{\sqrt{-2s_{j-1}}}(I - (s_{j-1} + s_j)(A + s_j I)^{-1})\mathbf{z}_{j-1}$.
   - 2.2. $\mathbf{Z}_j = [\mathbf{Z}_{j-1}, \mathbf{z}_j]$.
   - If converged, stop.

At each iteration, the recurrence in Algorithm 6 thus requires system solves with a fixed number of right-hand sides corresponding to the number of columns of C_1. As for the generation of the rational Krylov subspace, a new block of linear systems needs to be solved as the shift varies. For a very sparse A and a small number of precomputed shifts, one could consider factorizing each of the matrices $A + s_j I$ by means of a sparse solver, and then back solving at each ADI iteration. The feasibility of this procedure is clearly problem and architecture dependent.

In [28], [29] some key relations are used to show that the residual norm can be computed efficiently. More precisely, it holds that $A\mathbf{Z}_j\mathbf{Z}_j^* + \mathbf{Z}_j\mathbf{Z}_j^* A^* + C_1 C_1^* = W_j W_j^*$, where W_j is a matrix of rank p (the rank of C_1) defined as (here with real poles)

$$W_j := W_{j-1} - 2s_j Q_j, \qquad W_0 = C_1,$$

where $Q_1 = (A + s_1)^{-1} C_1$ and $Q_j = (I - (s_j + s_{j-1})(A + s_j I)^{-1}) Q_{j-1}$, $j \geq 2$. This way,

$$\|A\mathbf{Z}_j\mathbf{Z}_j^* + \mathbf{Z}_j\mathbf{Z}_j^* A^* + C_1 C_1^*\|_\star = \|W_j W_j^*\|_\star = \|W_j^* W_j\|_\star,$$

$\star = 2, F$, where the last norm is cheap to compute and memory efficient.

Other recent contributions are devoted to further improving the computational cost per iteration. A strategy for reducing the number of solves was proposed under the name of "modified" low-rank Smith method in [112]. The idea is to compute the singular value decomposition (SVD) of the iterate at each step and, given a dropping tolerance, to replace the iterate with its best low-rank approximation. A main aspect is that the SVD is not recomputed from scratch; instead, it is updated after each step to include the new information and then truncated to retain only those singular values that lie above the specified tolerance. The use of the SVD exploits the fact that if $Z \approx V\Sigma U^*$ is a truncated SVD of Z, then $\mathbf{X} = ZZ^* \approx V\Sigma^2 V^*$ is the truncated spectral decomposition of $\mathbf{X}$, so that the low-rank factor can be readily maintained. In general, the procedure reduces the number of system solves per iteration in a way that depends on the linear independence of the new iterate columns with respect to those of previous steps. Since $\mathbf{X}$ belongs to a rational Krylov subspace, the SVD computation determines an orthogonal basis—the columns of V associated with numerically nonzero singular values—for the generated rational space. This fact makes the truncated ADI method even closer to projection methods based on the rational Krylov space: the only difference is the way the reduced solution matrix is computed; see [76] for a formalization of this relation by means of the skeleton approximation.

A bound for the difference between the traces of the solution $\mathbf{X}$ of the Lyapunov equation and its ADI approximation is proposed in [244], which shows that the right-hand side of the Lyapunov equation can sometimes greatly influence the eigenvalue decay rate of the solution.

Computation of the Shifts. The selection of the ADI parameters and their number has been a major topic of research for many years, since the performance of the method, in terms of number of iterations, heavily depends on those parameters.

Let A be stable. Assuming a zero starting approximate solution, from the general ADI recurrence it follows that the error matrix associated with the ADI approximation $\mathbf{X}_k^{\mathrm{ADI}}$ after k full iterations is given by (see also [198])

$$\mathbf{X} - \mathbf{X}_k^{\mathrm{ADI}} = (\bar{r}_k(A) r_k(-A)^{-1})\mathbf{X}\bar{r}_k(A)^* r_k(-A)^{-*}, \tag{53}$$

$$r_k(z) = \prod_{i=1}^{k}(s_i - z), \qquad \bar{r}_k(z) = \prod_{i=1}^{k}(\bar{s}_i - z).$$

This expression shows that for A normal, for a fixed k optimal parameters can be obtained by solving the minimax problem

$$\min_{s_1,\ldots,s_k \in \mathbb{C}^-} \max_{\lambda \in \Lambda(A)} \prod_{j=1}^{k} \left| \frac{\lambda - s_j}{\lambda + s_j} \right|. \tag{54}$$

The value of k is adjusted so that the set $\{s_1, \ldots, s_k\}$ is closed under conjugation in the case that A is real. It is worth mentioning that it can be computationally advantageous to repeatedly apply a modest number of poles, rather than use a larger set of poles that gives a marginally faster convergence rate for the scalar rational approximation problem, if the cost of applying these poles is significant. For instance, ad hoc implementations may consider applying the same pole multiple times in a row, so that the costly factorization of the shifted matrix is only performed periodically. For A with real spectrum, the minimax problem in (54) was solved by Zolotaryov with the discrete point set replaced by a real finite interval; if A is also symmetric, this leads to an asymptotically optimal linear convergence rate for the approximation. The optimal parameters are then given as (see, e.g., [81])

$$s_j = \mathrm{dn}\left[\frac{(2j-1)K}{2k}, m\right], \quad j = 1, \ldots, k,$$

where dn is a Jacobian elliptic function, and K is the complete elliptic integral of the first kind, of modulus m [1]. Generalizations to the case when the complex spectrum lies in certain specified complex regions Ω were discussed in [81]. However, it was only with the heuristic approach of Penzl in [198] that the computation of *suboptimal* ADI parameters became a more manageable procedure. The proposed strategy is performed as a preprocessing of the actual ADI computation: consider the Krylov subspaces $K_{k_A}(A, c)$, $K_{k_{A^{-1}}}(A^{-1}, c)$ for some vector c, and let V, W be such that their orthonormal columns span the two spaces, respectively. Let Ω_+, Ω_- be the regions containing the eigenvalues of V^*AV and of W^*AW (the Ritz values). The key idea in [198] is to replace the spectrum of A with the region $\Omega := \Omega_+ \cup \Omega_-$ and then solve the minimax problem (54) in Ω. The set Ω may be regarded as a reasonable approximation to the region of interest, the convex hull of the spectrum of A, and it can be more cheaply computed, especially for small $k_A, k_{A^{-1}}$; see [198] and the package [200] for more technical details. An adverse effect of this preprocessing is its computational cost: for rank-one C the cost induced by the generation of both $K_{k_A}(A, c)$, $K_{k_{A^{-1}}}(A^{-1}, c)$ for some vector c to determine the suboptimal poles is comparable to that of, e.g., the construction of the extended Krylov subspace of corresponding dimension; however, by the time good suboptimal poles are determined, the extended Krylov approach has also computed an approximate solution to the Lyapunov equation.

In spite of the significant improvements in the ADI parameter estimation, however, the method remains quite sensitive to the choice of these shifts, and performance can vary dramatically even for small changes in $k_A, k_{A^{-1}}$; see, e.g., the experiments in [226]. Adaptive strategies for pole selection such as those derived for the rational Krylov subspace in [77] are hard to obtain, since a basis for the generated space is not readily available. Nonetheless, these considerations have led to the investigation of hybrid approaches, which are described later in this section.

It was observed in [174] that the ADI method actually generates a (block) rational Krylov subspace for the given vector of shifts $\mathbf{s}_k = [s_1, \ldots, s_k]$. The connection between the ADI method and the Galerkin method with the rational Krylov subspace $\mathcal{K}_k(A, C_1, \mathbf{s}_k) = \mathrm{range}([(A+s_1 I)^{-1}C_1, \ldots, (A+s_k I)^{-1}C_1])$ can be made more precise when the two methods are used with the same parameters.

THEOREM 11 (see [76, Thm. 3.4]). *Assume that the field of values of A and s_j, $j = 1, \ldots, k$, lie in the same half complex plane, and that C_1 has rank one. Let the columns of V form an orthonormal basis of $\mathcal{K}_k(A, C_1, \mathbf{s}_k)$, and let λ_j, $j = 1, \ldots, k$, be the Ritz values of A onto $\mathcal{K}_k(A, C_1, \mathbf{s}_k)$, that is, λ_j are the eigenvalues of V^*AV. Then the ADI approximation coincides with the Galerkin approximate solution $\mathbf{X}_k$ with*

$\mathcal{K}_k(A, C_1, \mathbf{s}_k)$ *if and only if* $s_j = \overline{\lambda_j}$, $j = 1, \ldots, k$ *(under a suitable index permutation for the* λ_j*'s).*

The condition $s_j = \overline{\lambda_j}$, $j = 1, \ldots, k$, is seldom satisfied when the shifts are obtained by either an adaptive procedure or a Penzl-style preprocessing (however, see [111] for an iterative process that approximates such a set of parameters in the context of optimal model order reduction). We also recall that the bound of Proposition 5 shows that ADI cannot give much better results than the Galerkin approach with the rational space, while it is known that for poor poles ADI may give much larger residuals than in the optimal case [20]. Moreover, the lack of some form of optimality condition, e.g., orthogonality, seems to penalize the ADI idea; this problem was explored in recent work summarized in the next paragraph. Selected numerical experiments comparing the adaptive rational Krylov subspace method and ADI can be found in [76].

Hybrid ADI Methods. It was observed in [37] that

> The most criticized property of the ADI iteration for solving Lyapunov equations is its demand for a set of good shift parameters to ensure fast convergence. [...] Most of the [computationally cheap parameters] are suboptimal in many cases and thus fast convergence can indeed not be guaranteed. Additionally, if the convergence is slow, the low-rank Cholesky factors may grow without adding essential information in subsequent iteration steps.

In [37] it was thus suggested to combine the subspace projection idea with the ADI recurrence. The projection is performed onto the space spanned by the columns of the current ADI factor, the idea being motivated by the fact the ADI solution factor belongs to the rational Krylov subspace with the same shifts as ADI. The projection is performed every $\hat{k}$ ADI iterations by computing an orthonormal basis spanning the range of the current factor, and the small size projected equation is solved by means of a Schur-type method (see section 5.1). Since the number of columns grows at each iteration, the cost of computing the orthonormal basis significantly increases. To overcome this problem, the authors suggest truncating the obtained projected solution so that a small-rank factor is retained for the next ADI iteration. More technical implementation details can be found in [37]. The idea is very reminiscent of a restarting process in the standard linear system framework, although here the truncation is performed in a different fashion. To complete the parallel with linear system solves, this procedure may be viewed as a hybrid restarted process, where a rational function (here the ADI single step) is applied to the solution before restart; see, e.g., [230] for a review of polynomial acceleration procedures of restarted methods in the linear system setting. The resulting process is called the Galerkin projection accelerated LRCF-ADI (LRCF-ADI-GP). Note that although ADI does not require that A be either positive or negative definite, the extra projection step is ensured to not break down only under the additional definiteness constraint. It is also interesting to observe that, without the truncation of the projected solution, the procedure might be mathematically equivalent to the Galerkin method in the rational Krylov subspace obtained with the same shift parameters; a formal proof still needs to be carried out. Selected numerical experiments comparing the adaptive *tangential* rational Krylov subspace method and projected ADI can be found in [78].

We also mention the procedure proposed in [139], where the continuous Lyapunov equation is first transformed into a discrete (Smith) Lyapunov equation with rational matrix functions as coefficient matrices, and is then solved by means of the global Krylov subspace method. This may be viewed as a preconditioning strategy.

5.2.3. Spectral, Sparse Format, and Other Methods. As for the Sylvester equation, the Kronecker formulation can be used to restate the matrix equation as the very large linear system

$$\mathcal{A}\mathbf{x} := (I_n \otimes A + A^* \otimes I_n)\mathbf{x} = c, \qquad \mathbf{x} = \mathrm{vec}(\mathbf{X}),\ c = \mathrm{vec}(C), \tag{55}$$

of size n^2, where n is the size of A; see, e.g., [126] for an early attempt to solve the system by exploiting the structure of $\mathcal{A}$. For A symmetric and positive definite, the convergence rate of CG applied to the Kronecker formulation is influenced by the condition number $\kappa(\mathcal{A}) = \kappa(A)$, whereas the convergence rate of the Galerkin procedure directly applied to the original Lyapunov equation is influenced by $\kappa(A + \lambda_{\min} I)$ (see Theorem 9), which can be significantly smaller than $\kappa(A)$. This analysis justifies the better performance of projection methods applied to the matrix equation. A second possibly stronger argument is given by memory requirements: the Kronecker formulation requires n^2-length vectors. Nonetheless, it was recently shown in [185] that when solving (55) floating point operations can be carried out so as to lower memory storage from $\mathcal{O}(n^2)$ to $\mathcal{O}(n)$. Moreover, a standard Krylov subspace method for (55) can take full advantage of the structure, since matrix-vector multiplications can be rewritten as matrix-matrix operations.

A possible way to overcome slow convergence is to choose an effective preconditioning strategy that can improve the spectral properties of the coefficient matrix $\mathcal{A}$. Hochbruck and Starke used a Krylov subspace solver for the system (55), and they investigated SSOR and ADI iteration (with a fixed number of iterations) as operator-based preconditioners; see also [185] for some implementation aspects of preconditioning strategies. More recently, a flexible GMRES approach was proposed in [44], which allowed for a variable ADI preconditioning step. Very preliminary numerical results report promising performance of the Kronecker formulation, while taking into account the matrix structure. These approaches may have broader applications for more general matrix equations; see the discussion in section 7.2.

A rather different approach consists of using an appropriately modified version of the sign function iteration depicted in (45). As memory requirements are excessive in its original form for large scale problems, two major amendments have been explored (see, e.g., [16]): (i) a sparsified version of A, so as to substantially reduce the computation and storage of Z_k^{-1}; (ii) for $C = C_1 C_1^*$, a factored version of the approximation $\widetilde{\mathbf{X}}$, so that only a tall factor need be iterated. The latter problem was addressed in [33], where the following coupled iteration was proposed:

$$A_0 = A, \quad B_0 = C_1, A_{k+1} = \frac{1}{2}(A_k + A_k^{-1}), \quad B_{k+1} = \frac{1}{\sqrt{2}}[B_k,\ A_k^{-1}B_k],\ k = 0, 1, \ldots,$$

giving $\mathbf{Y} = \frac{1}{\sqrt{2}} \lim_{k\to\infty} B_k$, with $\mathbf{Y}\mathbf{Y}^* = \mathbf{X}$. Note that the number of columns of B_k is doubled at each iteration, therefore a rank reduction is suggested in [33]. A recent extensive investigation of the performance of this type of approach can be found in [224]; the discussion in [224] in fact addresses the generalized Sylvester equation.

Item (i), namely, reducing the cost of dealing with the explicit inverse of large matrices, may be addressed by exploiting data sparse matrix representations and approximate arithmetic. In [18], but also in previous related works for the algebraic

Riccati equation, the $\mathcal{H}$-matrix format was used (see section 4.4.3). If $\mathrm{Inv}_{\mathcal{H}}(A)$ denotes the inverse in the $\mathcal{H}$-matrix format, then the coupled recurrence above can be performed as

$$A_{k+1} = \frac{1}{2}(A_k + \mathrm{Inv}_{\mathcal{H}}(A_k)), \quad B_{k+1} = \frac{1}{\sqrt{2}}[B_k,\ \mathrm{Inv}_{\mathcal{H}}(A_k)B_k],\ k = 0, 1, \ldots,$$

where the sum to obtain A_{k+1} is intended in $\mathcal{H}$-matrix format. More implementation details can be found in [18]. According to the analysis performed there, the error induced by the new format can be controlled while performing the rank reduction of B_{k+1}, so that the format errors do not grow unboundedly with k; these results are in agreement with the general theory of $\mathcal{H}$-matrices for Riccati equations developed in [105]. In [18], the derivation with the $\mathcal{H}$-matrix format is extended to the case of the generalized Lyapunov equation (see section 7). Numerical experiments show that the $\mathcal{H}$-format allows the sign function iteration to be employed for medium size problems ($\mathcal{O}(10000)$), for which the dense algorithm requires excessive memory allocation. Finally, an example comparing a linear multigrid solver using $\mathcal{H}$-format matrices with ADI is reported in [103, sec. 7.6], showing that for that specific example the multigrid approach is about ten times faster than ADI (implemented in the same framework), although both methods scale linearly with the number of multigrid levels.

We conclude this section with strategies that are more explicitly based on invariant subspaces. All methods considered assume that the maximum rank of a sufficiently accurate approximate solution is either known or given. Therefore, the context in which these approaches are used is different from that of previous methods.

The integral representation of $\mathbf{X}$ in (22) and the decay of the singular values of $\mathbf{X}$ suggest various eigenvalue-based strategies. One such method focuses on approximating the leading invariant subspace of $\mathbf{X}$. In [127] and [128] an approximate power iteration (API) approach was proposed, which aims to approximate the dominant eigenvectors of $\mathbf{X}$. The method is closely related to the power iteration and the Lanczos method for computing the extremal eigenpairs of a positive definite symmetric matrix, and the authors report good convergence properties when the eigenvalues associated with the sought-after eigenvectors are away from zero and well separated from the others, so that a good low-rank approximation of $\mathbf{X}$ can be determined. The method works under the assumption that A is negative definite, as with projection methods. The API method applies the power method to $\mathbf{X}$, which is only known implicitly and approximately by means of products of the type $\mathbf{Y} = \mathbf{X}v$ through the solution of the auxiliary "tall" Sylvester equation

$$A\mathbf{Y} + \mathbf{Y}\Theta + q = 0, \tag{56}$$

where $\Theta = v^*Av$ is a small square matrix and $q = C_1C_1^*v$ (see section 4.3). The numerical experiments reported in [128] for small problems seem to imply that API is a promising method for the approximation of the leading eigenvectors of $\mathbf{X}$, without the computation of $\mathbf{X}$ itself. The approach is reminiscent of the implicitly restarted Arnoldi method [171], although each iteration requires the solution of a Sylvester equation. A variant of this approach was proposed in [252] to overcome misconvergence caused by the omission of the term $\mathbf{X}(I - vv^*)A^*v$ in (56). Motivated by [128], an algorithm combining the power method and (implicitly) the ADI iteration was proposed in [190]; see [189] for a more thorough presentation of these approaches.

With the same aim of approximating the leading invariant subspace of $\mathbf{X}$ of given dimension, the procedure explored in [109] performs a refined numerical approximation by repeatedly integrating the dynamical system associated with the Lyapunov equation as the basis for an orthogonal power iteration.

A somewhat related approach was proposed in [231], which exploits the popular *proper orthogonal decomposition* (POD) approach employed in reduced order modeling of large scale dynamical systems [32]. The idea is to collect a sample of m approximate solutions to a sequence of associated linear time-dependent differential equations with different starting data and, for a chosen k, form a rank-k approximate Lyapunov solution. The approach relies on the integral representation of the Lyapunov solution and, according to the author, it is particularly appropriate for infinite-dimensional problems.

Finally, a novel and very different approach was recently proposed by Vandereycken and Vandewalle in [250] for A symmetric and positive definite: the method finds a low-rank approximation to $\mathbf{X}$ by minimizing the function

$$f : \mathcal{M}_k \to \mathbb{R}, \quad X \mapsto \mathrm{trace}(XAX) - \mathrm{trace}(XC)$$

on the manifold $\mathcal{M}_k$ of symmetric and positive semidefinite matrices of rank k in $\mathbb{R}^{n\times n}$, namely,

$$\min_{X\in\mathcal{M}_k} f(X).$$

When $\mathbf{X}_\star$ is the true solution to the Lyapunov equation, it was proved in [250] that $\|\mathrm{vec}(X-\mathbf{X}_\star)\|^2_{\mathcal{A}} = 2f(X)+2\mathrm{trace}(\mathbf{X}_\star A\mathbf{X}_\star)$ for all $X \in \mathcal{M}_k$, with $\mathcal{A}$ as in (55), so that the minimization of f corresponds to the minimization of the error in the energy norm, which is defined as $\|x\|^2_{\mathcal{A}} = x^*\mathcal{A}x$. By using the smoothness of $\mathcal{M}_k$ the problem is solved within a Riemann optimization framework, which allows one to embed the rank constraint in the space and solve an unconstrained minimization problem by means of a Riemann trust-region method, a second-order model based on the Hessian [2]. At convergence of the minimization process, if the current solution rank is not sufficiently accurate, the process is restarted basically from scratch. As a result, the method may be appealing when the optimal rank is approximately known a priori; otherwise, the approach may not be competitive with respect to other strategies discussed so far.

6. The Stein and Discrete Lyapunov Equations. The Stein and the discrete Sylvester equations are the discrete-time counterpart of the (continuous-time) equations discussed in the previous sections, and they naturally stem from a discrete-time system; see (5) and, e.g., [4, sec. 4.3]. Other relevant applications include, for instance, statistics [152], [151], probability [10], and spectral analysis [133]. These equations are also a computational tool in the design of control systems [156], and in the coprime matrix fraction description of linear systems [269].

The Stein equation may be written as

$$\mathbf{X} + A\mathbf{X}B = C, \tag{57}$$

where it is assumed that the eigenvalues of A and B are contained in the open unit disk. The discrete-time Lyapunov equation is obtained by choosing $B = -A^*$, in which case, if C is symmetric and if a solution $\mathbf{X}$ exists, then $\mathbf{X}$ has to be symmetric. In the context of inertia theory, for $C \succeq 0$ the discrete-time Lyapunov equation allows one to analyze the proximity of $\mathrm{spec}(A)$ to the unit circle and the proximity of $\mathrm{spec}(\mathbf{X})$ to the imaginary axis; see, [168, sec. 13.2] and also, e.g., [260], [172] for more specialized results.

Under the condition that $\lambda_i(A)\lambda_j(B) \neq -1$ for all i, j, the solution $\mathbf{X}$ exists and is unique for any C (see, e.g., [167]), and this is highlighted by the Kronecker form of

(57), given as $(I + B^* \otimes A)\mathbf{x} = c$, where $\mathbf{x} = \text{vec}(\mathbf{X})$ and $c = \text{vec}(C)$. Necessary and sufficient conditions for the existence and uniqueness of the solution $\mathbf{X}$ were obtained in [259] as a generalization of the property (19) for the Sylvester equation. Inertia and other transformation-based results for $B = -A^*$ can be derived in a natural manner from those for the Lyapunov equation; see, e.g., [221], [222]. We also refer the reader to [166] for a solution expressed in terms of the companion form of the given matrices, and to [38] for related computational considerations.

To numerically solve the equation for, say, B nonsingular, one could work with $\mathbf{X}B^{-1} + A\mathbf{X} = CB^{-1}$, which is a standard Sylvester equation, and then adopt one of the solution methods from previous sections. In fact, (57) is nothing more than a generalized Sylvester equation as in (1) with special choices of the first two coefficient matrices. For large B, the matrix B^{-1} should not be formed explicitly, but instead its action used within iterative methods.

Forming B^{-1} explicitly is also not recommended in the small size case, whenever B is ill-conditioned. Alternative transformations that bring the discrete equation to standard form are given by (for $B = -A^*$; see [204])

$$\tilde{A}\tilde{\mathbf{X}} + \tilde{\mathbf{X}}\tilde{A}^* = C, \quad \text{with } \tilde{A} = (A - I)^{-1}(A + I),\ \mathbf{X} = \frac{1}{2}(\tilde{A} - I)^*\tilde{\mathbf{X}}(\tilde{A} - I),$$

and (see [12], [201], [148])

$$\tilde{A}\mathbf{X} + \mathbf{X}\tilde{A}^* = \tilde{C}, \quad \text{with } \tilde{A} = (A - I)(A + I)^{-1}, \tilde{C} = 2(A^* + I)^{-1}C(A + I)^{-1},$$

where it is assumed that the inversions are well defined. In general, however, the same stability considerations as for methods using A^{-1} apply.

All these difficulties encourage solving the discrete equations (57) directly. A Schur-form type method for small size coefficient matrices that directly deals with (57) can be found, e.g., in [13], while a generalization of the "continuous-time" Hammarling method was proposed by Hammarling himself in [115].

In [251], Varga established a rank-two updating formula for the Cholesky factors in Hammarling's algorithm for solving the real, nonnegative definite Stein equation. As a generalization of his algorithm for the Lyapunov equation, a block variant of Hammarling's method for the discrete-time Lyapunov equation is suggested in [158].

In spite of the strong similarity with the standard equation, directly attacking (57) is an interesting problem in its own right, especially for A and B of large dimensions and with either of the two matrices singular. For a low-rank C, projection methods are applicable to solve (57) and an approximate solution $\widetilde{\mathbf{X}} = V_k\mathbf{Y}W_k^*$ can be determined, where the columns of V_k and W_k span approximation spaces associated with A and B^*, respectively. For instance, a global Krylov subspace approach was proposed in [138, sec. 5], [140], and its implementation is a natural modification of that used for the standard Sylvester equation. Similar derivations can be obtained for other Krylov-based methods.

The discrete-time Lyapunov equation motivated the development of the Smith method [233], which is at the basis of the modern ADI iteration for the Lyapunov equation. For A d-stable (i.e., with eigenvalues inside the unit circle), the unique solution to (57) with $B = -A^*$ can be written as $\mathbf{X} = \sum_{j=0}^{\infty} A^j C (A^j)^*$, and it is real symmetric and positive semidefinite if C is. The (convergent) Smith iteration is defined as

$$\mathbf{X}_0 = 0, \quad \mathbf{X}_{k+1} = C + A\mathbf{X}_k A^*,$$

with a closed form given by $\mathbf{X}_k = \sum_{j=1}^{k} A^{j-1}C(A^{j-1})^*$. Faster (quadratic) convergence can be achieved with the *squared* Smith method, which becomes of interest in the large scale case precisely for C of small rank [198]. The iteration is generically given as

$$\mathbf{X} = A^{2^{k+1}}\mathbf{X}(A^{2^{k+1}})^* + \sum_{i=0}^{2^{k+1}-1} A^i C (A^i)^*, \quad \mathbf{X} = \lim_{k\to\infty} \sum_{i=0}^{2^{k+1}-1} A^i C (A^i)^*.$$

The resulting recursion is given by $H_{k+1} = H_k + A_k H_k A_k^*$, $H_0 = C$, where $A_{k+1} = A_k^2$, so that $C_k \to \mathbf{X}$ as $k \to \infty$. By exploiting the low rank of $C = C_0 C_0^*$, $H_{k+1} = C_{k+1} C_{k+1}^*$ with $C_{k+1} = [C_k, A_k C_k]$. Therefore, the number of columns of C_{k+1} doubles at each iteration, and C_{k+1} is contained in a block Krylov subspace generated by A and C_0. Recent advances to make this recurrence more effective in terms of both computational costs and memory requirements include compressions, truncations, and restarts, with a tricky use of the underlying Krylov subspace [175], [216], [25]. In these references, estimates for the residual and error norms are also derived. Finally, we point out an ADI acceleration strategy in [216] (for $B = -A^*$) and in [25], which significantly improves the convergence speed. In fact, a major breakthrough for the Smith method consisted in combining its recurrence with the ADI idea, as developed in [198].

All these approaches rely on the fact that often the solution $\mathbf{X}$ has (numerical) rank much lower than n; indeed, in [25] it is shown for the Stein equation that if the eigenvalues of A and B lie inside the open unit disk and C has rank p,

$$\frac{\sigma_{kp+1}(\mathbf{X})}{\sigma_1(\mathbf{X})} \le \|A^k\|\|B^k\|,$$

indicating that the solution rank might indeed be small if the powers of A and B decrease rapidly in norm. In [216] the following estimate was derived for $B = -A^*$ and $\|A\| < 1$:

$$\frac{\sigma_{kp+1}(\mathbf{X})}{\sigma_1(\mathbf{X})} \le \frac{\|A\|^{2k}}{1 - \|A\|^2}.$$

In general, a computational comparison of various variants of the approaches based on the Smith iteration is still lacking, though highly desirable.

A related matrix equation is the $\top$-Stein equation, given by $\mathbf{X} = A\mathbf{X}^\top B + C$, whose solvability conditions have been recently analyzed in [177]. More generally, a broader class of matrix equations can be written as $\mathbf{X} = Af(\mathbf{X})B + C$, where $f(X) = X^\top$, $f(X) = \bar{X}$, or $f(X) = X^*$, whose analysis and numerical solution can be recast in terms of the Stein matrix equation [271]. This and more general forms of linear equations are discussed in the next section.

7. Generalized Linear Equations.

7.1. The Generalized Sylvester and Lyapunov Equations.

The term *generalized* refers to a very wide class of equations, which includes systems of matrix equations, bilinear equations, and problems where the coefficient matrices are rectangular. We start with the most common form of the generalized Sylvester equation, namely,

$$A\mathbf{X}D + E\mathbf{X}B = C, \tag{58}$$

which differs from (18) in the occurrence of coefficient matrices on both sides of the unknown solution $\mathbf{X}$.

If D and E are both nonsingular, left multiplication by E^{-1} and right multiplication by D^{-1} lead to a standard Sylvester equation, with the same solution matrix $\mathbf{X}$. If either E or D is ill-conditioned, such a transformation may lead to severe instabilities. This problem is common to other *generalized* equations we will encounter later in this section, and it justifies the development of solution methods that stick to the original form (58). The case of singular D and E, especially for $D = E^*$ and $B = A^*$, has an important role in the solution of differential-algebraic equations and descriptor systems [164]. The solution of (58) for E and D singular requires knowledge of the spectral projectors onto the right and left deflating subspaces of the *stable* pencils $\lambda E - A$ and $\lambda D - B$, associated with the finite eigenvalues, along with the right and left deflating subspaces associated with the eigenvalue at infinity. In such a setting, the right-hand side matrix is also projected onto the corresponding deflating subspaces and the equation is called the *projected* Sylvester equation. The numerical treatment of this matrix equation necessitates ad hoc procedures that appropriately and stably take into account the Weierstrass canonical form of the pencils $\lambda E - A$, $\lambda D - B$, from which the spectral projectors can be derived; we refer the reader to, e.g., [239], [178] and their references for further details on projected Sylvester equations.

The following result ensures the existence of a unique solution $\mathbf{X}$ to (58).

THEOREM 12 (see [58]). *The matrix equation $A\mathbf{X}D + E\mathbf{X}B = C$ has a unique solution if and only if*
(i) *the pairs (A, E) and $(D, -B)$ are regular pencils;*
(ii) *the spectra of (A, E) and $(B, -D)$ are disjoint.*[15]

Under the hypotheses of Theorem 12, uniqueness is thus still ensured if one of the matrices A, B, D, or E is singular, as long as the corresponding pencil is nonsingular.

A natural extension of the Bartels–Stewart method can be implemented for numerically solving (58) when dimensions are small, and this was discussed in [93], [94], [196], where the starting point is a QZ decomposition of the pencils (A, E) and (B, D) followed by the solution of a sequence of small (1-by-1 or 2-by-2) generalized Sylvester equations, which is performed using their Kronecker form. For C positive semidefinite and (A, E) stable, in [196] a generalization of the Hammarling method is also proposed. The algorithm developed in [93], [94] is also able to treat the case in which some specifically selected coefficient matrices are singular.

The large scale setting does not significantly differ from previous cases, as long as E, D are not too ill-conditioned. The problem can be recast as a standard Sylvester equation in $E^{-1}A$ and BD^{-1}. In the case of rational Krylov subspace and ADI methods, shifted systems can be solved with the coefficient matrix $(E^{-1}A + sI) = E^{-1}(A + sE)$, and analogously for systems with BD^{-1}. In the case of ill-conditioned E, D, one could consider using a specifically selected $\alpha \in \mathbb{R}$ (or $\alpha \in \mathbb{C}$) such that the two matrices $E + \alpha A$ and $D - \alpha B$ are better conditioned and the solution uniqueness is ensured, and rewrite (58) as the equivalent generalized Sylvester matrix equation $A\mathbf{X}(D - \alpha B) + (E + \alpha A)\mathbf{X}B = C$.

We mention the specific application of global Krylov subspace methods (see section 4.4), which are obtained by using the mapping $\mathcal{M}(X) = AXD + EXB$; therefore, they can be applied in general to the equation $\sum_{i=1}^{q} A_i X B_i = C$, as is done in [46]. Note that this kind of approach can only be applied to medium size problems, as the matrix formulation involves dense matrices. We recall once again that there is a tight

[15] Here the notion of disjoint spectra [58, formula (7)] should be understood in light of the definition of "spectral set" in generalized eigenvalue problems, as defined, for instance, in [238, Def. VI.1.1].

relation between global methods and the Kronecker form, which provides a good basis for the theoretical understanding of global methods.

A unique solution to the generalized Lyapunov equation

$$A\mathbf{X}E^* + E\mathbf{X}A^* = C \tag{59}$$

is ensured if and only if Theorem 12 applies, that is, all eigenvalues of the pencil (A, E) are finite and they do not have pairwise zero sum. As a consequence, a unique solution is only obtained if one of the matrices A, E is nonsingular. In this case one can recast (6) as a standard Lyapunov equation.

To avoid stability problems caused by a possibly ill-conditioned E or A, it is usually preferred to work with E and A implicitly. This is realized by performing a simultaneous Schur decomposition of E and A, $E = QSZ^*$ and $A = QTZ^*$, with S and T (complex) upper triangular [186]. Plugging in this transformation, (59) becomes $QTZ^*\mathbf{X}ZS^*Q^* + QSZ^*\mathbf{X}Z^*T^*Q^* = C$, that is,

$$T\widehat{\mathbf{X}}S^* + S\widehat{\mathbf{X}}T^* = Q^*CQ, \quad \widehat{\mathbf{X}} = Z^*\mathbf{X}Z.$$

The elements of $\widehat{\mathbf{X}}$ can then be obtained by exploiting the structure of T and S [116].

A different approach adapts the matrix sign function iteration in (46) to this more general context, and it is shown in [33] that it is applicable under the hypothesis that the Lyapunov equation is stable. In the case of C in factorized form in (59), a recurrence is proposed in [33] to generate an approximation to the Cholesky-type factor of the resulting semidefinite solution $\mathbf{X}$. Comparisons in terms of memory requirements and floating point operations with respect to the generalized Hammarling method (see [196]) are also reported in [33]. We also refer the reader to [196] for some estimates of the separation[16] and the condition number of the operator associated with (59), which is important to assessing the accuracy of the computed solution.

7.2. Bilinear, Constrained, and Other Linear Equations. Other generalizations of the Sylvester equation have attracted the attention of many researchers. In some cases the standard procedure for their solution consists in solving a (sequence of) related standard Sylvester equation(s), so that the computational core is the numerical solution of the latter by means of some of the procedures discussed in previous sections. We thus list here some of the possible generalizations more often encountered and employed in real applications. We start by considering the case when the two coefficient matrices can be rectangular. This gives the equation

$$A\mathbf{X} + \mathbf{Y}B = C, \tag{60}$$

where $\mathbf{X}$, $\mathbf{Y}$ are both unknown, and A, B, and C are all rectangular matrices of conforming dimensions. Equations of this type arise in control theory, for instance, in output regulation with internal stability, where the matrices are in fact polynomial matrices (see, e.g., [263] and references therein). The following theorem is a first result on the existence and uniqueness of the pair $\mathbf{X}$, $\mathbf{Y}$ and is reported as originally stated in [211]; see also more recent advanced developments in [83].

THEOREM 13 (see [211]). *The necessary and sufficient condition that the equation $A\mathbf{X} - \mathbf{Y}B = C$, where A, B, and C are $m \times r$, $s \times n$, and $m \times n$ matrices, respectively,*

[16] Defined as $\mathrm{sep}_p(A, E) = \min_{\|X\|_p=1} \|A^*XE + E^*XA\|_p$, with $p = 2, F$.

with elements in a field $\mathcal{F}$, has a solution $\mathbf{X}, \mathbf{Y}$ of order $r \times n$ and $m \times s$, respectively, and with elements in $\mathcal{F}$, is that the matrices

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

be equivalent.[17]

The matrix equivalence in the theorem can be explicitly obtained as

$$\begin{bmatrix} I & \mathbf{Y} \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & -\mathbf{X} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},$$

where $\mathbf{Y}$ and $\mathbf{X}$ are the solution matrices to the matrix equation (60). Solvability conditions of this equation can be derived from [238, Chap. VI]; the result is stated next.

PROPOSITION 14 (see [238]). *Assume A and B are nonsingular. Then the problem $A\mathbf{X} + \mathbf{Y}B = C$ has a unique solution when the spectra of A and B are disjoint.*

These are extensions of the analogous result for the standard Sylvester equation; see (19) and [211], [131]. Note that by setting, for instance, $\mathbf{U} = [\mathbf{X}, \mathbf{Y}]$, (60) can be rewritten as

$$A\mathbf{U} \begin{bmatrix} I_n \\ 0 \end{bmatrix} + \mathbf{U} \begin{bmatrix} 0 \\ B \end{bmatrix} = C$$

in the single unknown matrix $\mathbf{U}$ [157].

The two-sided version of (60) is given by

$$A\mathbf{X}D + E\mathbf{Y}B = C,$$

and this is an example of the more complex bilinear equations with several left-hand side terms considered in the literature; see, e.g., [269] and references therein.

A typical generalization is given by the bilinear equation

$$A\mathbf{X}D + E\mathbf{X}B = C\mathbf{Y} + F, \tag{61}$$

where the pair $(\mathbf{X}, \mathbf{Y})$ is to be determined, and $\mathbf{X}$ occurs in two different terms. Theoretical aspects are collected in [266] and also in [267], where closed forms for $(\mathbf{X}, \mathbf{Y})$ are given. In [268] general parametric expressions for the solution matrices $\mathbf{X}$ and $\mathbf{Y}$ are also obtained, under the hypothesis that D is full rank and F is the zero matrix.

The main objective in the aforementioned papers is in fact the solution of *systems* of bilinear matrix equations

$$\begin{cases} A_1\mathbf{X} + \mathbf{Y}B_1 = C_1, \\ A_2\mathbf{X} + \mathbf{Y}B_2 = C_2 \end{cases} \tag{62}$$

(see, e.g., [147], [79]), for which a recent perturbation analysis can be found in [180]. These systems can arise, for instance, in the numerical treatment of systems of stochastic partial differential equations, giving rise to large and sparse coefficient matrices; see, e.g., (16) and [85]. The system (62) is an important step in deflating subspace

[17] Two $n \times m$ matrices A and B are called equivalent if $B = Q^{-1}AP$ for some invertible matrices P and Q of matching dimensions.

computations for pencils [145], [238, Chap. VI], [58]. Indeed, the system can be formulated in terms of a transformation $\mathcal{P}^{-1}(\mathcal{M} - \lambda\mathcal{N})\mathcal{Q}$ onto a block diagonal form of the matrix pencil

$$\mathcal{M} - \lambda\mathcal{N} = \begin{bmatrix} A_1 & -C_1 \\ 0 & -B_1 \end{bmatrix} - \lambda \begin{bmatrix} A_2 & -C_2 \\ 0 & -B_2 \end{bmatrix}.$$

The pair $(\mathbf{X}, \mathbf{Y})$ is sought such that

$$\mathcal{P}^{-1}(\mathcal{M} - \lambda\mathcal{N})\mathcal{Q} := \begin{bmatrix} I & -\mathbf{Y} \\ 0 & I \end{bmatrix} (\mathcal{M} - \lambda\mathcal{N}) \begin{bmatrix} I & \mathbf{X} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & -B_1 \end{bmatrix} - \lambda \begin{bmatrix} A_2 & 0 \\ 0 & -B_2 \end{bmatrix}.$$

Setting

$$\mathcal{L} : (\mathbf{X}, \mathbf{Y}) \mapsto (A_1\mathbf{X} + \mathbf{Y}B_1, A_2\mathbf{X} + \mathbf{Y}B_2),$$

the problem of solvability of (62) corresponds to that of determining when the operator $\mathcal{L}$ is nonsingular. It turns out that for (A_i, B_i), $i = 1, 2$ regular pairs, $\mathcal{L}$ is nonsingular if and only if the spectra of (A_1, B_1) and (A_2, B_2) are disjoint [238, Thm. VI.1.11]; this result also leads to the one in Proposition 14. The operator $\mathcal{L}$ is a generalization of the corresponding operator for the Sylvester equation; its sensitivity in terms of distance to singularity can be defined analogously to the sep operator; see, e.g., [147], [238, sec. VI.2.4].

From a numerical standpoint, the most reliable approach for small scale computations was proposed in [147] and further developed in [146] and is based on the stable generalized Schur method, which applies the QZ algorithm to the pairs of coefficient matrices; a perturbation analysis is also included. Few other alternatives have yet been explored that go beyond a cleverly implemented Kronecker formulation. The idea suggested in [79] amounts to "expanding" the two equations into a single equation of larger size, whose solution contains both $\mathbf{X}$ and $\mathbf{Y}$, but requires the Jordan decomposition of some of the coefficient matrices. A similar framework is used in [73] where more than two unknown matrices are allowed, and an approximate solution is obtained by means of a least squares approach. It is not clear how any of these procedures can be adapted to the large scale setting.

The number of linear matrix equations and unknown matrices can in fact be quite large, as is discussed, for instance, in [45]. Necessary and sufficient conditions for the resulting systems to have a solution pair are studied in [257]. Computationally speaking, this general case has so far only been treated by using the Kronecker formulation, so that only very small problems have been tackled; however, see [269], where the problem of solving the set of matrix equations is recast as an optimization problem.

A special class of *nonlinear* problems is given by the following Sylvester-observer equation, which stems from the problem of determining a *reduced*-order observer model [248], [182]. Find matrices $\mathbf{X}, \mathbf{Y}$, and $\mathbf{Z}$ such that

$$\mathbf{X}A - \mathbf{Y}\mathbf{X} = \mathbf{Z}C, \quad \begin{bmatrix} \mathbf{X} \\ C \end{bmatrix} \text{ invertible}, \tag{63}$$

where A and C are known matrices with C having few rows. A solution to (63) exists for any choice of spectrum of $\mathbf{Y}$, and therefore this spectrum can be predetermined; a choice that makes $\mathbf{Y}$ a stable matrix also ensures convergence of the reduced-order observer; we refer the reader to [182] for more details on these aspects. A possible way to address the solution of (63) is to choose $\mathbf{Y}$ and $\mathbf{Z}$ arbitrarily and then solve the resulting Sylvester equation for $\mathbf{X}$. Early approaches in this direction did not lead to a numerically stable method. For small size matrices, the reduction to Hessenberg form proposed by Van Dooren in [248] is still one of the most effective methods for solving (63). The algorithm is based on a reduction to "staircase form" of the pair (A, C) and on the determination of the solution $\mathbf{X}$ with a particular structure in a different coordinate system. We also refer the reader to [249] for a more detailed survey on methods for dense matrices. More recently, other approaches have been proposed: for instance, a block generalization of the method in [248] was proposed in [54]; moreover, in [55] the authors proposed a block algorithm for determining a full-rank solution, and which seems to be most appropriate for large scale problems with sparse A. In this latter setting, a successful early method was proposed in [67]. The approach first assumes that $\mathbf{Z}C$ is rank one and then exploits the resemblance between the Sylvester-observer equation and the Arnoldi relation (34). As a by-product of the method, the authors in [67] also derive an algorithm for solving the partial pole assignment problem for large and sparse A, which is generalized in [66] to higher rank of $\mathbf{Z}C$. The authors in [52] propose a new strategy for a priori choosing the eigenvalues of $\mathbf{Y}$ that makes the algorithm in [67] more efficient. From a control theory point of view, the possibility of determining a reduced-order model is also important in the derivation of stable closed-loop systems, giving rise to a well-exercised eigenvalue assignment problem. We refer the reader to, e.g., [64] for a brief survey on this and other related problems.

Within the Sylvester-observer equation, we can formulate the problem in a slightly different manner, namely, by means of a constraint (see, e.g., [241], [245], [187]), and it can be stated as follows (see, e.g., [11]): *Given* $A \in \mathbb{R}^{n\times n}$ *stable,* $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{m\times n}$, *and* $F \in \mathbb{R}^{(n-m)\times(n-m)}$, *find* $\mathbf{Y} \in \mathbb{R}^{(n-m)\times m}$ *and* $\mathbf{X} \in \mathbb{R}^{(n-m)\times n}$ *such that*

$$\mathbf{X}A - F\mathbf{X} = \mathbf{Y}C, \quad \textit{with} \quad \mathbf{X}B = 0, \tag{64}$$

so that $[\mathbf{X}; C]$ *is full rank.*

The problem may be viewed as a homogeneous system of linear matrix equations, a generalization of (62), with two terms in $\mathbf{X}$ as in (61) and $C_1 = 0$, $C_2 = 0$; however, there is no need to expand it by means of the Kronecker product. In [11] the authors provide necessary and sufficient conditions for a solution to (64) to exist, and propose an algorithm for its computation in the small scale case. The main ingredients are a QR factorization and the solution of a standard Sylvester equation. In [95] a modification of this method was presented to handle the case of almost singular Sylvester equations. The large scale case has been recently addressed in [220] by generalizing the method in [11]. We also point to [56] for an approach that handles a "regional pole-placement constraint" on F in (64) for a descriptor system, and to [63] and its references for further theoretical properties.

Going back to a single unknown matrix, other "generalized" Lyapunov equations more in the spirit of (2) include extra linear terms,

$$A\mathbf{X} + \mathbf{X}A^* + N\mathbf{X}N^* + C_1C_1^* = 0, \tag{65}$$

and they stem, for instance, from complex dynamical systems like the one in (10). We refer the reader to [107] for sufficient conditions on the existence of the controllability and observability Gramians; more complex forms involve more structured matrices N; see, e.g., [22], [59]. In fact, more terms of the type $N_j\mathbf{X}N_j^*$, $j = 1, 2, \ldots$, could arise to fulfill more general model requests. Polynomial and infinite-dimensional systems are also of interest; see, e.g., [264] and [61], respectively, and their references. In addition to a robust Kronecker-form based iteration reviewed in [62, secs. 3.1–4], Damm in [62] proposed a regular splitting for the numerical solution of (65), yielding the iterative scheme

$$A\mathbf{X}_{k+1} + \mathbf{X}_{k+1}A^* = -N\mathbf{X}_k N^* - C_1 C_1^*, \quad \mathbf{X}_0 = 0,$$

which entails the solution of a sequence of standard Lyapunov equations. Convergence to $\mathbf{X}$ is obtained if the spectrum of A is sufficiently far away from the imaginary axis. In [62, sec. 4] the generalized case of the Lyapunov operator is also treated. In the recent article [23] a thorough discussion and contextualization of the algebraic problem in stochastic model order reduction can be found. In [21], various methods for the Lyapunov equation, such as ADI and projection techniques, are adapted to the setting of (65), including sparse format approaches for the Kronecker formulation; reported experimental results on large problems seem to favor this sparse format approach, with the caveat that the sparsity and accuracy parameters must be tuned, as described in section 5.2.3.

An approach that may be appropriate for large scale problems is implicitly suggested in [8]. In the context of model order reduction, the following approximation space is introduced:

$$\text{range}(V) = \text{span}\left\{ \bigcup_{k=1}^{r} \text{range}\{V^{(k)}\} \right\}, \tag{66}$$

with $\text{range}(V^{(1)}) := K_q(A^{-1}, A^{-1}C_1)$ and

$$\text{range}(V^{(k)}) := K_q(A^{-1}, A^{-1}NV^{(k-1)}), k = 2, \ldots, r.$$

Using a Galerkin approximation onto range(V), (65) can be reduced and solved with a direct procedure; a possible implementation of this idea was recently proposed in [21]. Another approach for solving multilinear systems in Kronecker form was analyzed in [163], in which a tensor-based form for the approximate solution is considered. Such a strategy is well suited to the approximation of *parameterized* linear systems, which arise, for instance, in certain discretization strategies for the numerical solution of stochastic partial differential equations [6]. Data sparse methods associated with the Kronecker formulation may provide a possibly successful avenue for attacking the general linear multiterm matrix equation (2); to the best of our knowledge, no attempts have been made in this direction so far for really large problems.

7.3. Sylvester-like and Lyapunov-like Equations. Sylvester- and Lyapunov-like linear matrix equations of the form [47], [39]

$$B\mathbf{X} + f(\mathbf{X})A = C, \quad A^*\mathbf{X} + f(\mathbf{X})A = C, \quad B, A, \mathbf{X} \in \mathbb{C}^{m\times n}, \tag{67}$$

with $f(X) = \bar{X}$, $f(X) = X^\top$, and $f(X) = X^*$, or their "discrete-time" variants (see section 6) are less common, but see, for instance, [161] for an occurrence in structured eigenvalue computation. The homogeneous case ($C = 0$) has been recently analyzed in [72], where a complete description of the solution in terms of the Kronecker canonical form of $A + \lambda f(B)$ is derived whenever information on this latter pencil is available. These equations have attracted increasing interest in the past few years, with recent contributions on the necessary and sufficient conditions for their solvability, for any right-hand side matrix C [134]; a different proof of this result that also induces a numerical method is proposed in [253]. As an example of this type of result, in [51, Lem. 5.10] for $f(X) = X^\top$, it is proved that a unique solution $\mathbf{X}$ exists if and only if the pencil $A - \lambda B^\top$ is regular and if its spectrum is $\top$-reciprocal free,[18] with possibly the only exception of the unit eigenvalue, which should be simple.

Interestingly, it was recently shown that for A and B nonsingular, the problem $A\mathbf{X} + \mathbf{X}^\top B = C$ can be recast as a standard Sylvester equation. The following result is proved in [74].

PROPOSITION 15. *Assume that A and B are nonsingular. If $\mathbf{X}$ is a solution to the matrix equation $A\mathbf{X}+\mathbf{X}^\top B = C$, then $\mathbf{X}$ is also a solution to the Sylvester matrix equation*

$$(B^{-\top}A)\mathbf{X} - \mathbf{X}(A^{-\top}B) = B^{-\top}C - B^{-\top}CA^{-\top}B. \tag{68}$$

The reverse also holds if (68) *admits a unique solution* $\mathbf{X}$.

Under the given hypotheses, this result allows one to solve the $\top$-Sylvester equation by means of procedures for the standard Sylvester equation, with the caveat of maintaining good stability properties of the problem.

Going back to the original formulation in (67), in [70] an algorithm that relies on the generalized Schur decomposition of the pair $(A, f(B))$ (via the QZ algorithm) is proposed to determine $\mathbf{X}$ for small A and B. For $f(X) = X^\top$ this can be briefly summarized as follows:

(1) Decompose $A = URV$ and $B^\top = USV$, with U, V unitary and R, S upper triangular.

(2) Compute $E = VCV^\top$.

(3) Solve $S^\top W + W^\top R = E$ element by element.

(4) Form $X = \bar{U}W\bar{V}$.

The solution of the equation in step (3) is also treated in detail in [70].

Recent developments have considered the case where both A and B have large dimensions. In particular, in [74] projection methods are derived to solve the $\top$-Sylvester equation for the case when A and B are nonsingular. They generate right and left approximation spaces $\mathcal{V}_m$ and $\mathcal{W}_m$, respectively, satisfying $B^\top\mathcal{V}_m = \mathcal{W}_m$, so that a suitable Petrov–Galerkin condition can be imposed. The reduction yields a small Sylvester-like equation of the same form, which can be solved with the Schur decomposition strategy above. The two approximation spaces can be chosen as any of the Krylov-based spaces described in previous sections; we refer the reader to [74] for algorithmic details.

In [265] a closed-form solution to the equation for $f(X) = \bar{X}$ is considered, together with the set of all possible solutions for (61) and for the bilinear problem $A\bar{\mathbf{X}} + B\mathbf{Y} = \mathbf{X}F$. Chiang, Duan, Feng, Wu, and their collaborators have thoroughly investigated these formulations and their role in control applications.

A particular extension of this class of problems is given by polynomial equations. Consider a polynomial matrix $R(\xi) = R_0 + R_1\xi + \cdots + R_\ell\xi^\ell$ in the unknown ξ, where R_i are constant square matrices such that $\det(R(\xi))$ is not identically zero, and let Z be a square polynomial matrix satisfying $Z(\xi) = Z(-\xi)^\top$. The equation

$$R(-\xi)^\top\mathbf{X}(\xi) + \mathbf{X}(-\xi)^\top R(\xi) = Z(\xi)$$

in the square polynomial matrix $\mathbf{X}$ is called the polynomial Sylvester equation. This special equation plays a role in the computation of integrals of quadratic functions of

[18] A set of complex numbers $\{\lambda_1, \ldots, \lambda_k\}$ is $\star$-*reciprocal free* if $\lambda_i \neq 1/\lambda_j^\star$ for any $1 \leq i, j \leq k$. Typically, $\star = \top$ or $\star = *$, so that $\lambda_j^\star$ is λ_j or $\bar{\lambda}_j$, respectively.

the variables of a system and their derivatives (see, e.g., [242]), and in the stability theory for high-order differential equations. In [195] the authors focus on the case when the right-hand side has the form $Z(\xi) = Q(-\xi)^\top \Sigma Q(\xi)$, where Q is a real rectangular polynomial matrix in ξ such that QR^{-1} is a matrix of strictly proper rational functions and Σ is a diagonal, signature matrix. An iterative solution method inspired by the method of Faddeev for the computation of the matrix resolvents is also described; we refer the reader to [117] for a detailed derivation of the Faddeev sequence in connection with the solution of Lyapunov and Sylvester equations. More general equations include polynomial Diophantine matrix equations in the form $D(\xi)\mathbf{X}(\xi) + N(\xi)\mathbf{Y}(\xi) = F(\xi)$; in [156] closed-form solutions are presented, which could be used to numerically solve small size equations. In the large scale setting, this problem is computationally unsolved.

Finally, special attention should be paid to the homogeneous version of the Sylvester-like equation previously discussed with $B = A$,

$$A\mathbf{X}^\top + \mathbf{X}A = 0. \tag{69}$$

For each fixed complex matrix A, the solution space to this latter equation is a Lie algebra equipped with Lie bracket $[X, Y] := XY - YX$. We refer the reader to the recent articles [71], [92] and their references for more details.

8. Software and High Performance Computation. Reliable software for solving matrix equations has been available for a long time, due to its fundamental role in control applications; in particular, the SLICE Library was made available in 1986. Early in the 1990s the SLICOT library (http://www.slicot.org/ [246]) replaced SLICE, and since then a large number of additions and improvements have been included; see, e.g., [232], [27]. Most recent versions of MATLAB[19] also rely on calls to SLICOT routines within the control-related toolboxes. SLICOT includes a large variety of codes for model reduction and nonlinear problems on sequential and parallel architectures; as a workhorse, both the Bartels–Stewart algorithm and the Hessenberg–Schur algorithm are implemented. The Bartels–Stewart algorithm for triangular matrices is also included in LAPACK. Functions for solving the Lyapunov equation are also available in other computational environments, such as *Mathematica*.[20] Related projects have led to the development of additional codes, which are usually available either on more general websites or directly from the authors. For instance, specifically oriented to linear matrix equations in the MATLAB framework, the `lyapack` set of routines developed by Penzl in [200] has been particularly successful as a possible implementation of the ADI method for large scale Lyapunov equations, relying on preprocessing for the computation of quasi-optimal parameters. These routines were included in the NICONET project repository (http://www.icm.tu-bs.de/NICONET/). The MESS set[21] by Saak, Mena, and Benner is the successor to the `lyapack` package, with the aim of fully exploiting the capabilities of newer releases of MATLAB. In addition, MESS allows for the solution of a larger variety of matrix equations associated with the differential Riccati equation. A rather detailed list of routines for solving control-related matrix equations is provided in the book by Sima [223] and in the more recent book by Datta [65].

[19] MATLAB is a registered trademark of The MathWorks Inc.

[20] *Mathematica* is a registered trademark of Wolfram Research.

[21] Available at http://www.en.mpi-magdeburg.mpg.de/mpcsc/mitarbeiter/saak/Software/mess.php?lang=en.

A number of benchmark problems have been made available for testing purposes. In addition to those available in the NICONET website, a variety of datasets is available in the Oberwolfach collection[22] accompanied by a well-documented description of the originating application problems; see also the description in [160].

Refined implementations of structured linear equation methods have been proposed for high performance computations. In particular, the efficient solution of triangular and quasi-triangular Sylvester equations has been discussed in [206], [202]. A high performance library for triangular Sylvester-type matrix equations (continuous- and discrete-time) is also available at http://www8.cs.umu.se/~isak/recsy/, while a parallel SCALAPACK-style version of this software, called SCASY, is available at http://www8.cs.umu.se/~granat/scasy.html. Some of the SLICOT routines are overloaded in these libraries; see [143], [144], [100] for more information on their implementation on parallel architectures.

In [129] an early parallel algorithm was developed to solve medium size ($0 < n \leq 1000$) Lyapunov problems with a banded and negative definite matrix A; experiments with a shared memory multiprocessor machine (Alliant FX-8) can also be found. The approach is similar in spirit to classical iterative linear system methods such as Jacobi and Gauss–Seidel. More recently, specialized parallel algorithms for Lyapunov, Stein, and other generalized matrix equations for different modern architectures have been presented by a number of authors; see, e.g., [205] for the Cray T3E, [34], [35] employing a cluster of PCs, and [24] for hybrid CPU-GPU platforms. The use of approaches based either on the square Smith iteration or on iterative techniques for the matrix sign function, as opposed to the Schur decomposition, is key to obtaining good parallel performance.

Systems of matrix equations were implemented in a parallel environment in [45] and references therein. A parallel algorithm for the small scale solution to the multi-input Sylvester-observer equation (see section 7.2) was proposed in [43] and tested on two shared-memory vector machines.

9. Concluding Remarks and Future Outlook.

The solution of linear matrix equations has always attracted the attention of the engineering and scientific communities. The reliability of efficient core numerical linear algebra methods has made the solution of these matrix equations increasingly popular in application problem modeling. A good understanding of the theoretical tools and of the variety of numerical methods available for Sylvester-type equations provides a solid ground for attacking more general—nonlinear, multiterm, or multifunctional—matrix equations. In particular, the efficient solution of multiterm matrix equations such as those in (2) represents the next frontier for numerical linear algebra, as it is currently one of the major bottlenecks in the numerical treatment of PDEs involving stochastic terms; see section 3. Advances in this direction will be tightly related to those being made in the solution of linear systems with tensor product structure, which in the simplest case can be written as

$$\mathcal{A}\mathbf{x} = b \quad \text{with} \quad \mathcal{A} = \sum_{j=1}^{k} I_{n_1} \otimes \cdots \otimes I_{n_{j-1}} \otimes A_j \otimes I_{n_{j+1}} \cdots \otimes I_{n_k}. \tag{70}$$

This problem is a further level of generalization of the standard Sylvester equation, where the solution is a k-way tensor whose size explodes with k even for modest values

[22] Available at http://portal.uni-freiburg.de/imteksimulation/downloads/benchmark.

of n_j [154]. The complex Kronecker structure arising in (70) makes the problem very hard to even analyze, and its size calls for truncation or reduction procedures that rely on approximation theory and hierarchical data structures; see, e.g., [106], [191], [192]. Scientific computing applications dealing with many variables can exploit these data tools to considerably lower the computational complexity of their model; see, e.g., [149] for a recent survey. Among the very recent projection methods used for the solution of (70), we find Krylov subspace based procedures that considerably generalize methods used for the two-dimensional case; see, e.g., [162], [9]. We envisage that a lot of scientific research will be devoted to multiterm and multidimensional problems in forthcoming years.

We have mainly limited our presentation to linear problems. Nonlinear matrix equations have a crucial and ever increasing role in many applications: for instance, the popular algebraic Riccati equation (see [167]) has a leading role in control applications and is an important tool in eigenvalue problems; we refer the reader to [42] for a very recent presentation of the rich literature on computational methods. Other fully nonlinear equations include, e.g, matrix eigenvalue problems [82], [184] and equations of the type $\mathbf{X} + A^\top F(\mathbf{X})A = Q$, where F is a properly defined nonlinear function of $\mathbf{X}$ (see, e.g., [207] and references therein), together with matrix equations involving powers of $\mathbf{X}$.

Linear matrix equations with special properties arise when dealing with *periodic* dynamical systems. These problems give rise to periodic counterparts of the equations we have analyzed, such as Lyapunov and Sylvester equations. Corresponding Schur forms can be used for their solution, and necessary and sufficient conditions for a periodic discrete-time system to be equivalent to a time-invariant system are known; thorough treatments with developments on both the theoretical and algorithmic fronts, mainly on small size problems, have been carried out by Byers, Van Dooren, Sreedhar, Varga, and their collaborators.

Acknowledgments. We are grateful to three reviewers for their numerous helpful remarks. We would like to thank Ulrike Baur, Peter Benner, Michele Benzi, Lars Grasedyck, Yiding Lin, Daniel B. Szyld, and Paul Van Dooren for helpful comments and pointers to many references. We are indebted to Daniel Szyld for providing us with [129]. We are grateful to Daniel Kressner and Froilán M. Dopico for carefully reading and amending a previous version of this article, and for several helpful references. We are also grateful to the librarians of the Math Library at the Dipartimento di Matematica, Università di Bologna for their professional dedication.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965. (Cited on pp. 395, 414)

[2] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008. (Cited on p. 418)

[3] M. I. Ahmad, I. Jaimoukha, and M. Frangos, *Krylov subspace restart scheme for solving large-scale Sylvester equations*, in 2010 American Control Conference, Marriott Waterfront, Baltimore, MD, 2010, pp. 5726–5731. (Cited on p. 401)

[4] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, Adv. Des. Control 6, SIAM, Philadelphia, 2005. (Cited on pp. 381, 382, 395, 400, 418)

[5] A. C. Antoulas, D. C. Sorensen, and Y. Zhou, *On the decay rate of Hankel singular values and related issues*, Systems Control Lett., 46 (2002), pp. 323–342. (Cited on p. 405)

[6] I. Babuška, R. Tempone, and G. E. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Numer. Anal., 42 (2004), pp. 800–825. (Cited on p. 426)

[7] Z. Bai and J. Demmel, *Design of a parallel nonsymmetric eigenroutine toolbox, Part* I, in Proceedings of the 6th SIAM Conference on Parallel Processing for Scientific Computing, R. F. Sincovec et al., ed., 1993, pp. 391–398 (Cited on p. 407)

[8] Z. Bai and D. Skoogh, *A projection method for model reduction of bilinear dynamical systems*, Linear Algebra Appl., 415 (2006), pp. 406–425. (Cited on p. 426)

[9] J. Ballani and L. Grasedyck, *A projection method to solve linear systems in tensor format*, Numer. Linear Algebra Appl., 19 (2012), pp. 27–43. (Cited on p. 430)

[10] A. D. Barbour and S. Utev, *Solving the Stein equation in compound Poisson approximation*, Adv. Appl. Probab., 30 (1998), pp. 449–475. (Cited on pp. 384, 418)

[11] J. B. Barlow, M. M. Monahemi, and D. P. O'Leary, *Constrained matrix Sylvester equations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1–9. (Cited on p. 425)

[12] S. Barnett, *Simplification of the Lyapunov matrix equation* $A^TPA-P=-Q$, IEEE Trans. Automat. Control, 19 (1974), pp. 446–447. (Cited on p. 419)

[13] A. Y. Barraud, *A numerical algorithm to solve* $A^TXA-X=Q$, IEEE Trans. Automat. Control, AC-22 (1977), pp. 883–885. (Cited on p. 419)

[14] J. F. Barrett, *Bibliography of A.M. Lyapunov's work*, Internat. J. Control, 55 (1992), pp. 785–790. (Cited on p. 378)

[15] R. H. Bartels and G. W. Stewart, *Algorithm* 432: *Solution of the Matrix Equation* $AX+XB=C$, Comm. ACM, 15 (1972), pp. 820–826. (Cited on pp. 379, 390)

[16] U. Baur, *Control-Oriented Model Reduction for Parabolic Systems*, Ph.D. thesis, Institut für Mathematik, TU Berlin, 2008. (Cited on p. 416)

[17] U. Baur, *Low rank solution of data-sparse Sylvester equations*, Numer. Linear Algebra Appl., 15 (2008), pp. 837–851. (Cited on p. 404)

[18] U. Baur and P. Benner, *Factorized solution of Lyapunov equations based on hierarchical matrix arithmetic*, Computing, 78 (2006), pp. 211–234. (Cited on pp. 406, 416, 417)

[19] U. Baur and P. Benner, *Cross-Gramian based model reduction for data-sparse systems*, Electron. Trans. Numer. Anal., 31 (2008), pp. 256–270. (Cited on p. 404)

[20] B. Beckermann, *An error analysis for rational Galerkin projection applied to the Sylvester equation*, SIAM J. Numer. Anal., 49 (2011), pp. 2430–2450. (Cited on pp. 401, 403, 409, 415)

[21] P. Benner and T. Breiten, *Low rank methods for a class of generalized Lyapunov equations and related issues*, Numer. Math., 124 (2013), pp. 441–470. (Cited on p. 426)

[22] P. Benner, T. Breiten, and T. Damm, *Generalised tangential interpolation for model reduction of discrete-time MIMO bilinear systems*, Internat. J. Control, 84 (2011), pp. 1398–1407. (Cited on p. 425)

[23] P. Benner and T. Damm, *Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems*, SIAM J. Control Optim., 49 (2011), pp. 686–711. (Cited on pp. 385, 426)

[24] P. Benner, P. Ezzatti, D. Kressner, E. S. Quintana-Ortí, and A. Rem'on, *A mixed-precision algorithm for the solution of Lyapunov equations on hybrid CPU-GPU platforms*, Parallel Comput., 37 (2011), pp. 439–450. (Cited on p. 429)

[25] P. Benner, G. El Khoury, and M. Sadkane, *On the squared Smith method for large-scale Stein equations*, Numer. Linear Algebra Appl., 21 (2014), pp. 645–665. (Cited on p. 420)

[26] P. Benner, M. Køhler, and J. Saak, *Sparse-Dense Sylvester Equations in* $\mathcal{H}_2$*-Model Order Reduction*, Tech. Report MPIMD/11-11, Max Planck Institute Magdeburg Preprint, 2011. Available from http://www.mpi-magdeburg.mpg.de/preprints/. (Cited on p. 392)

[27] P. Benner, D. Kressner, V. Sima, and A. Varga, *The SLICOT Toolboxes: A Survey*, SLICOT Working Note 2009-1, www.slicot.org, 2009. (Cited on pp. 390, 428)

[28] P. Benner, P. Kürschner, and J. Saak, *Efficient handling of complex shift parameters in the low-rank Cholesky factor ADI method*, Numer. Algorithms, 62 (2013), pp. 225–251. (Cited on p. 412)

[29] P. Benner, P. Kürschner, and J. Saak, *A reformulated low-rank ADI iteration with explicit residual factors*, PAMM Proc. Appl. Math. Mech., 13 (2013), pp. 585–588. (Cited on p. 412)

[30] P. Benner, J.-R. Li, and T. Penzl, *Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems*, Numer. Linear Algebra Appl., 15 (2008), pp. 1–23. (Cited on p. 412)

[31] P. Benner, R.-C. Li, and N. Truhar, *On the ADI method for Sylvester equations*, J. Comput. Appl. Math., 233 (2009), pp. 1035–1045. (Cited on p. 402)

[32] P. Benner, V. Mehrmann, and D. Sorensen, eds., *Dimension Reduction of Large-Scale Systems*, Lect. Notes Comput. Sci. Eng., Springer-Verlag, Berlin, Heidelberg, 2005. (Cited on pp. 381, 383, 418)

[33] P. Benner and E. S. Quintana-Ortí, *Solving stable generalized Lyapunov equations with the matrix sign function*, Numer. Algorithms, 20 (1999), pp. 75–100. (Cited on pp. 406, 416, 422)

[34] P. Benner, E. S. Quintana-Ortì, and G. Quintana-Ortì, *Solving stable Stein equations on distributed memory computers*, in EuroPar'99 Parallel Processing, P. Amestoy, P. Berger, M. Daydè, I. Duff, V. Frayssè, L. Giraud, and D. Ruiz, eds., Lecture Notes in Comput. Sci. 1685, Springer-Verlag, New York, 1999, pp. 1120–1123. (Cited on p. 429)

[35] P. Benner, E. S. Quintana-Ortì, and G. Quintana-Ortì, *Numerical solution of discrete stable linear matrix equations on multicomputers*, Parallel Algorithms Appl., 17 (2002), pp. 127–146. (Cited on p. 429)

[36] P. Benner, E. S. Quintana-Ortì, and G. Quintana-Ortì, *Solving stable Sylvester equations via rational iterative schemes*, J. Sci. Comput., 28 (2006), pp. 51–83. (Cited on p. 391)

[37] P. Benner and J. Saak, *A Galerkin-Newton-ADI Method for Solving Large-Scale Algebraic Riccati Equations*, Tech. Report SPP1253-090, Deutsche Forschungsgemeinschaft - Priority Program 1253, 2010. (Cited on p. 415)

[38] A. Betser, N. Cohen, and E. Zeheb, *On solving the Lyapunov and Stein equations for a companion matrix*, Systems Control Lett., 25 (1995), pp. 211–218. (Cited on pp. 406, 419)

[39] J. H. Bevis, F. J. Hall, and R. E. Hartwig, *The matrix equation $A\bar{X} - XB = C$ and its special cases*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 348–359. (Cited on p. 426)

[40] R. Bhatia and P. Rosenthal, *How and why to solve the operator equation $AX - XB = Y$*, Bull. London Math. Soc., 29 (1997), pp. 1–21. (Cited on p. 379)

[41] W. G. Bickley and J. McNamee, *Matrix and other direct methods for the solution of linear difference equation*, Philos. Trans. Roy. Soc. London Ser. A, 252 (1960), pp. 69–131. (Cited on pp. 385, 391)

[42] D. A. Bini, B. Iannazzo, and B. Meini, *Numerical Solution of Algebraic Riccati Equations*, SIAM, Philadelphia, 2012. (Cited on pp. 383, 384, 407, 430)

[43] C. H. Bischof, B. N. Datta, and A. Purkayastha, *A parallel algorithm for the multi-input Sylvester-observer equation*, in Proceedings 1992 American Control Conference, 1992, pp. 152–156. (Cited on p. 429)

[44] M. Bollhöfer and A. K. Eppler, *A structure preserving FGMRES method for solving large Lyapunov equations*, in Progress in Industrial Mathematics at ECMI 2010, Math. Ind. 17, M. Günther, A. Bartel, M. Brunk, S. Schoeps, and M. Striebel, eds., Springer, Berlin, 2012, pp. 131–136. (Cited on p. 416)

[45] I. Borno, *Parallel computation of the solutions of coupled algebraic Lyapunov equations*, Automatica, 31 (1995), pp. 1345–1347. (Cited on pp. 424, 429)

[46] A. Bouhamidi and K. Jbilou, *A note on the numerical approximate solutions for generalized Sylvester matrix equations with applications*, Appl. Math. Comput., 206 (2008), pp. 687–694. (Cited on pp. 395, 421)

[47] H. W. Braden, *The equations $A^T X \pm X^T A = B$*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 295–302. (Cited on p. 426)

[48] S. Brahma and B. Datta, *An optimization approach for minimum norm and robust partial quadratic eigenvalue assignment problems for vibrating structures*, J. Sound Vibration, 324 (2009), pp. 471–489. (Cited on p. 383)

[49] R. Byers, *A LINPACK-style condition estimator for the equation $AX + XB^T = C$*, IEEE Trans. Automat. Control, AC-29 (1984), pp. 926–928. (Cited on p. 389)

[50] R. Byers, *Solving the algebraic Riccati equation with the matrix sign function*, Linear Algebra Appl., 85 (1987), pp. 267–279. (Cited on p. 407)

[51] R. Byers and D. Kressner, *Structured condition numbers for invariant subspaces*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 326–347. (Cited on p. 427)

[52] D. Calvetti, B. Lewis, and L. Reichel, *On the solution of large Sylvester-observer equations*, Numer. Linear Algebra Appl., 8 (2001), pp. 435–451. (Cited on p. 425)

[53] 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. (Cited on p. 384)

[54] J. Carvalho and B. A. Datta, *A block algorithm for the Sylvester-observer equation arising in state-estimation*, in Proceedings IEEE International Conference on Decision and Control, Orlando, FL, 2001, pp. 3898–3903. (Cited on p. 425)

[55] J. Carvalho, K. Datta, and Y. Hong, *A new block algorithm for full-rank solution of the Sylvester-observer equation*, IEEE Trans. Automat. Control, 48 (2003), pp. 2223–2228. (Cited on p. 425)

[56] E. B. Castelan and V. G. da Silva, *On the solution of a Sylvester equation appearing in descriptor systems control theory*, Systems Control Lett., 54 (2005), pp. 109–117. (Cited on p. 425)

[57] C. Chen and D. Schonfeld, *Pose estimation from multiple cameras based on Sylvester's equation*, Computer Vision and Image Understanding, 114 (2010), pp. 652–666. (Cited on p. 384)

[58] K. E. Chu, *The solution of the matrix equations $AXB - CXD = E$ and $(YA - DZ, YC - BZ) = (E, F)$*, Linear Algebra Appl., 93 (1987), pp. 93–105. (Cited on pp. 421, 424)

[59] M. Condon and R. Ivanov, *Nonlinear systems - algebraic Gramians and model reduction*, Compel, 24 (2005), pp. 202–219. (Cited on p. 425)

[60] M. J. Corless and A. E. Frazho, *Linear Systems and Control: An Operator Perspective*, Pure Appl. Math., Marcel Dekker, New York, Basel, 2003. (Cited on p. 381)

[61] O. L. V. Costa and C. S. Kubrusly, *Lyapunov equation for infinite-dimensional discrete bilinear systems*, Systems Control Lett., 17 (1991), pp. 71–77. (Cited on p. 425)

[62] T. Damm, *Direct methods and ADI-preconditioned Krylov subspace methods for generalized Lyapunov equations*, Numer. Linear Algebra Appl., 15 (2008), pp. 853–871. (Cited on p. 426)

[63] M. Darouach, *Solution to Sylvester equation associated to linear descriptor systems*, Systems Control Lett., 55 (2006), pp. 835–838. (Cited on p. 425)

[64] B. N. Datta, *Linear and numerical linear algebra in control theory: Some research problems*, Linear Algebra Appl., 197–198 (1994), pp. 755–790. (Cited on pp. 381, 425)

[65] B. N. Datta, *Numerical Methods for Linear Control Systems: Design and Analysis*, Academic Press, Cambridge, MA, 2004. (Cited on p. 428)

[66] B. N. Datta and C. Hetti, *Generalized Arnoldi methods for the Sylvester-observer equation and the multi-input pole placement problem*, in Proceedings of the 36th IEEE Conference on Decision & Control, San Diego, CA, 1997, pp. 4379–4383. (Cited on p. 425)

[67] B. N. Datta and Y. Saad, *Arnoldi methods for large Sylvester-like observer matrix equations, and an associated algorithm for partial spectrum assignment*, Linear Algebra Appl., 154–156 (1991), pp. 225–244. (Cited on p. 425)

[68] J. Demmel, *Three methods for refining estimates of invariant subspaces*, Computing, 38 (1987), pp. 43–57. (Cited on p. 383)

[69] E. de Souza and S. P. Bhattacharyya, *Controllability, observability and the solution of $AX - XB = C$*, Linear Algebra Appl., 39 (1981), pp. 167–188. (Cited on p. 388)

[70] F. De Terán and F. M. Dopico, *Consistency and efficient solution of the Sylvester equation for $\star$-congruence*, Electron. J. Linear Algebra, 22 (2011), pp. 849–863. (Cited on p. 427)

[71] F. De Terán and F. M. Dopico, *The solution of the equation $XA + AX^T = 0$ and its application to the theory of orbits*, Linear Algebra Appl., 434 (2011), pp. 44–67. (Cited on p. 428)

[72] F. De Terán, F. M. Dopico, N. Guillery, D. Montealegre, and N. Reyes, *The solution of the equation $AX + X^\star B = 0$*, Linear Algebra Appl., 438 (2013), pp. 2817–2860. (Cited on p. 426)

[73] F. Ding and T. Chen, *Iterative least-squares solutions of coupled Sylvester matrix equations*, Systems Control Lett., 54 (2005), pp. 95–107. (Cited on p. 424)

[74] F. M. Dopico, J. González, D. Kressner, and V. Simoncini, *Projection methods for large-scale T-Sylvester equations*, Math. Comp., 2016, doi:10.1090/MCOM/3081. (Cited on p. 427)

[75] V. Druskin and L. Knizhnerman, *Extended Krylov subspaces: Approximation of the matrix square root and related functions*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 755–771. (Cited on p. 399)

[76] V. Druskin, L. Knizhnerman, and V. Simoncini, *Analysis of the rational Krylov subspace and ADI methods for solving the Lyapunov equation*, SIAM J. Numer. Anal., 49 (2011), pp. 1875–1898. (Cited on pp. 409, 413, 414, 415)

[77] V. Druskin and V. Simoncini, *Adaptive rational Krylov subspaces for large-scale dynamical systems*, Systems Control Lett., 60 (2011), pp. 546–560. (Cited on pp. 409, 410, 414)

[78] V. Druskin, V. Simoncini, and M. Zaslavsky, *Adaptive tangential interpolation in rational Krylov subspaces for MIMO dynamical systems*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 476–498. (Cited on pp. 411, 415)

[79] G. Duan, *A note on combined generalized Sylvester matrix equations*, J. Control Theory Appl., 4 (2004), pp. 397–400. (Cited on pp. 423, 424)

[80] N. S. Ellner and E. L. Wachspress, *New ADI model problem applications*, in Proceedings of 1986 ACM Fall Joint Computer Conference, Dallas, TX, IEEE Press, Los Alamitos, CA, 1986, pp. 528–534. (Cited on pp. 384, 395, 396, 401)

[81] N. S. Ellner and E. L. Wachspress, *Alternating direction implicit iteration for systems with complex spectra*, SIAM J. Numer. Anal., 28 (1991), pp. 859–870. (Cited on pp. 396, 414)

[82] H. C. Elman and M. Wu, *Lyapunov inverse iteration for computing a few rightmost eigenvalues of large generalized eigenvalue problems*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1685–1707. (Cited on pp. 383, 430)

[83] E. Emre and L. M. Silverman, *The equation $XR+QY = \Phi$: A characterization of solutions*, SIAM J. Control Optim., 19 (1981), pp. 33–38. (Cited on p. 422)

[84] M. Epton, *Methods for the solution of $ADXD - BXC = E$ and its application in the numerical solution of implicit ordinary differential equations*, BIT, 20 (1980), pp. 341–345. (Cited on p. 384)

[85] O. G. Ernst, C. E. Powell, D. J. Silvester, and E. Ullmann, *Efficient solvers for a linear stochastic Galerkin mixed formulation of diffusion problems with random data*, SIAM J. Sci. Comput., 31 (2009), pp. 1424–1447. (Cited on pp. 386, 423)

[86] K. V. Fernando and H. Nicholson, *On the structure of balanced and other principal representations of SISO systems*, IEEE Tran. Automat. Control, 28 (1983), pp. 228–231. (Cited on p. 382)

[87] K. V. Fernando and H. Nicholson, *On a fundamental property of the cross-Gramian matrix*, IEEE Trans. Circ. Syst., CAS-31 (1984), pp. 504–505. (Cited on p. 382)

[88] G. Flagg and S. Gugercin, *On the ADI method for the Sylvester equation and the optimal-$H2$ points*, Appl. Numer. Math., 64 (2013), pp. 50–58. (Cited on p. 403)

[89] D. Furnival, H. Elman, and C. Powell, *$H(div)$ preconditioning for a mixed finite element formulation of the stochastic diffusion problem*, Math. Comp., 79 (2010), pp. 733–760. (Cited on p. 386)

[90] Z. Gajić and M. Qureshi, *Lyapunov Matrix Equation in System Stability and Control*, Math. Sci. Engrg., Academic Press, San Diego, CA, 1995. (Cited on pp. 379, 381, 383, 384, 388, 404, 406)

[91] K. Gallivan, A. Vandendorpe, and P. Van Dooren, *Sylvester equations and projection-based model reduction*, J. Comput. Appl. Math., 162 (2004), pp. 213–229. (Cited on p. 382)

[92] S. R. Garcia and A. L. Shoemaker, *On the matrix equation $AX+XA^T = 0$*, Linear Algebra Appl., 438 (2013), pp. 2740–2746. (Cited on p. 428)

[93] J. D. Gardiner, A. J. Laub, J. J. Amato, and C. B. Moler, *Solution of the Sylvester matrix equation $AXB^{\mathrm{T}} + CXD^{\mathrm{T}} = E$*, ACM Trans. Math. Software, 18 (1992), pp. 223–231. (Cited on p. 421)

[94] J. D. Gardiner, M. R. Wette, A. J. Laub, J. J. Amato, and C. B. Moler, *Algorithm 705: A FORTRAN-77 software package for solving the Sylvester matrix equation $AXB^{\mathrm{T}} + CXD^{\mathrm{T}} = E$*, ACM Trans. Math. Software, 18 (1992), pp. 232–238. (Cited on p. 421)

[95] A. R. Ghavimi and A. J. Laub, *Numerical methods for nearly singular constrained matrix Sylvester equations*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 212–221. (Cited on p. 425)

[96] I. Gohberg, T. Kailath, and V. Olshevsky, *Fast Gaussian elimination with partial pivoting for matrices with displacement structure*, Math. Comp., 64 (1995), pp. 1557–1576. (Cited on p. 391)

[97] I. Gohberg, P. Lancaster, and L. Rodman, *Invariant Subspaces of Matrices with Applications*, J. Wiley & Sons, New York, 1986. (Cited on p. 387)

[98] G. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD, 1996. (Cited on pp. 389, 390)

[99] G. H. Golub, S. Nash, and C. Van Loan, *A Hessenberg-Schur method for the problem $AX + XB = C$*, IEEE Trans. Automat. Control, AC-24 (1979), pp. 909–913. (Cited on p. 390)

[100] R. Granat, *Parallel algorithms and condition estimators for standard and generalized triangular Sylvester-type matrix equations*, in PARA'06 - State of the Art in Scientific and Parallel Computing, Lecture Notes in Comput. Sci. 4699, Springer, New York, 2007, pp. 127–136. (Cited on p. 429)

[101] L. Grasedyck, *Theorie und Anwendungen Hierarchischer Matrizen*, Ph.D. thesis, Mathematisch-Naturwissenschaftliche Fakultät, Christian-Albrechts-Universität zu Kiel, Germany, 2001. (Cited on p. 404)

[102] L. Grasedyck, *Existence of a low rank or $\mathcal{H}$-matrix approximant to the solution of a Sylvester equation*, Numer. Linear Algebra Appl., 11 (2004), pp. 371–389. (Cited on p. 404)

[103] L. Grasedyck, *Nonlinear multigrid for the solution of large-scale Riccati equations in low-rank and $\mathcal{H}$-matrix format*, Numer. Linear Algebra Appl., 15 (2008), pp. 779–807. (Cited on p. 417)

[104] L. Grasedyck and W. Hackbusch, *A multigrid method to solve large scale Sylvester equations*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 870–894. (Cited on pp. 403, 404)

[105] L. Grasedyck, W. Hackbusch, and B. Khoromskij, *Solution of large scale algebraic matrix Riccati equations by use of hierarchical matrices*, Computing, 70 (2003), pp. 121–165. (Cited on p. 417)

[106] L. Grasedyck, D. Kressner, and C. Tobler, *A literature survey of low-rank tensor approximation techniques*, GAMM-Mitt., 36 (2013), pp. 53–78. (Cited on p. 430)

[107] W. S. Gray and J. Mesko, *Energy functions and algebraic Gramians for bilinear systems*, in Preprints of the 4th IFAC Nonlinear Control Systems Design Symposium 1998, Enschede, The Netherlands, Pergamon, Oxford, UK, 1998, pp. 101–106. (Cited on pp. 385, 425)

[108] E. Grimme, *Krylov Projection Methods for Model Reduction*, Ph.D. thesis, The University of Illinois at Urbana-Champaign, 1997. (Cited on p. 400)

[109] T. Gudmundsson and A. J. Laub, *Approximate solution of large sparse Lyapunov equations*, IEEE Trans. Automat. Control, 39 (1994), pp. 1110–1114. (Cited on p. 417)

[110] A. El Guennouni, K. Jbilou, and A. J. Riquet, *Block Krylov subspace methods for solving large Sylvester equations*, Numer. Algorithms, 29 (2002), pp. 75–96. (Cited on pp. 392, 394)

[111] S. Gugercin, A. C. Antoulas, and C. Beattie, *$\mathcal{H}_2$ model reduction for large-scale linear dynamical systems*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 609–638. (Cited on p. 415)

[112] S. Gugercin, D. C. Sorensen, and A. C. Antoulas, *A modified low-rank Smith method for large-scale Lyapunov equations*, Numer. Algorithms, 32 (2003), pp. 27–55. (Cited on p. 413)

[113] M. H. Gutknecht, *Krylov Subspace Algorithms for Systems with Multiple Right Hand Sides: A Collection of References*, 2004. Available at http://www.sam.math.ethz.ch/~mhg/talks/abstr/abstr.html. (Cited on pp. 380, 398)

[114] S. Güttel, *Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection*, GAMM Mitt., 36 (2013), pp. 8–31. (Cited on pp. 400, 401)

[115] S. Hammarling, *Numerical solution of the discrete-time, convergent, non-negative definite Lyapunov equation*, Systems Control Lett., 17 (1991), pp. 137–139. (Cited on p. 419)

[116] S. J. Hammarling, *Numerical solution of the stable, non-negative definite Lyapunov equation*, IMA J. Numer. Anal., 2 (1982), pp. 303–323. (Cited on pp. 406, 422)

[117] B. Hanzon and R. L. M. Peeters, *A Faddeev sequence method for solving Lyapunov and Sylvester equations*, Linear Algebra Appl., 241–243 (1996), pp. 401–430. (Cited on p. 428)

[118] C. Hartmann, B. Schäfer-Bung, and A. Zueva, *Balanced averaging of bilinear systems with applications to stochastic control*, SIAM J. Control Optim., 51 (2013), pp. 2356–2378. (Cited on p. 385)

[119] J. Z. Hearon, *Nonsingular solutions of $TA - BT = C$*, Linear Algebra Appl., 16 (1977), pp. 57–63. (Cited on p. 388)

[120] D. Heinig and U. Jungnickel, *Lyapunov equations for companion matrices*, Linear Algebra Appl., 76 (1986), pp. 137–147. (Cited on p. 406)

[121] G. Hewer and C. Kenney, *The sensitivity of the stable Lyapunov equation*, SIAM J. Control Optim., 26 (1988), pp. 321–344. (Cited on p. 405)

[122] M. Heyouni and K. Jbilou, *Matrix Krylov subspace methods for large scale model reduction problems*, Appl. Math. Comput., 181 (2006), pp. 1215–1228. (Cited on p. 411)

[123] N. J. Higham, *Perturbation theory and backward error for $AX - XB = C$*, BIT, 33 (1993), pp. 124–136. (Cited on pp. 389, 405)

[124] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, 1996. (Cited on p. 389)

[125] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008. (Cited on p. 401)

[126] M. Hochbruck and G. Starke, *Preconditioned Krylov subspace methods for Lyapunov matrix equations*, SIAM Matrix Anal. Appl., 16 (1995), pp. 156–171. (Cited on p. 416)

[127] A. S. Hodel and K. Poolla, *Numerical solution of very large, sparse Lyapunov equations through approximate power iteration*, in Proceedings of the 29th Conference on Decision and Control, IEEE, WA12, 1990, pp. 291–296. (Cited on p. 417)

[128] A. Hodel, B. Tenison, and K. Poolla, *Numerical solution of the Lyapunov equation by approximate power iteration*, Linear Algebra Appl., 236 (1996), pp. 205–230. (Cited on p. 417)

[129] A. S. Hodel, *Numerical Methods for the Solution of Large and Very Large, Sparse Lyapunov Equations*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1989. (Cited on pp. 407, 429, 430)

[130] A. S. Hodel and P. Misra, *Least-squares approximate solution of overdetermined Sylvester equations*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 279–290. (Cited on p. 397)

[131] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991. (Cited on pp. 379, 381, 387, 404, 423)
[132] D. Y. Hu and L. Reichel, *Krylov subspace methods for the Sylvester equation*, Linear Algebra Appl., 172 (1992), pp. 283–313. (Cited on pp. 397, 399, 401, 407, 411)
[133] G.-D. Hu and Q. Zhu, *Bounds of modulus of eigenvalues based on Stein equation*, Kybernetika, 46 (2010), pp. 655–664. (Cited on p. 418)
[134] Kh. D. Ikramov, *Conditions for unique solvability of the matrix equation* $AX + X^T B = C$, Dokl. Math., 81 (2010), pp. 63–65. (Cited on p. 426)
[135] I. M. Jaimoukha and E. M. Kasenally, *Oblique projection methods for large scale model reduction*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 602–627. (Cited on pp. 400, 411)
[136] I. M. Jaimoukha and E. M. Kasenally, *Implicitly restarted Krylov subspace methods for stable partial realizations*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 633–652. (Cited on p. 411)
[137] I. M. Jaimoukha and E. M. Kasenally, *Krylov subspace methods for solving large Lyapunov equations*, SIAM J. Numer. Anal., 31 (1994), pp. 227–251. (Cited on pp. 408, 411)
[138] K. Jbilou, *Low rank approximate solutions to large Sylvester matrix equations*, Appl. Math. Comput., 177 (2006), pp. 365–376. (Cited on pp. 399, 419)
[139] K. Jbilou, *ADI preconditioned Krylov methods for large Lyapunov matrix equations*, Linear Algebra Appl., 432 (2010), pp. 2473–2485. (Cited on p. 415)
[140] K. Jbilou and A. Messaoudi, *A computational method for symmetric Stein matrix equations*, in Numerical Linear Algebra in Signals, Systems and Control, Lecture Notes in Electrical Engrg. 80, Springer, New York, 2011, pp. 295–311. (Cited on p. 419)
[141] K. Jbilou, A. Messaoudi, and H. Sadok, *Global FOM and GMRES algorithms for matrix equations*, Appl. Numer. Math., 31 (1999), pp. 49–63. (Cited on p. 399)
[142] K. Jbilou and A. J. Riquet, *Projection methods for large Lyapunov matrix equations*, Linear Algebra Appl., 415 (2006), pp. 344–358. (Cited on p. 411)
[143] I. Jonsson and B. Kågström, *Recursive blocked algorithms for solving triangular systems, Part* I: *One-sided and coupled Sylvester-type matrix equations*, ACM Trans. Math. Software, 28 (2002), pp. 392–415. (Cited on pp. 390, 391, 429)
[144] I. Jonsson and B. Kågström, *Recursive blocked algorithms for solving triangular systems, Part* II: *Two-sided and generalized Sylvester and Lyapunov matrix equations*, ACM Trans. Math. Software, 28 (2002), pp. 416–435. (Cited on pp. 390, 429)
[145] B. Kagström and P. Van Dooren, *A generalized state-space approach for the additive decomposition of a transfer matrix*, Internat. J. Numer. Linear Algebra Appl., 1 (1992), pp. 165–181. (Cited on p. 424)
[146] B. Kågström and P. Poromaa, *LAPACK-style algorithms and software for solving the generalized Sylvester equation and estimating the separation between regular matrix pairs*, ACM Trans. Math. Software, 22 (1996), pp. 78–103; also as LAPACK Working Note 75. (Cited on p. 424)
[147] B. Kågström and L. Westin, *Generalized Schur methods with condition estimators for solving the generalized Sylvester equation*, IEEE Trans. Automat. Control, 34 (1989), pp. 745–751. (Cited on pp. 423, 424)
[148] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980. (Cited on p. 419)
[149] B. N. Khoromskij, *Tensors-structured numerical methods in scientific computing: Survey on recent advances*, Chemometrics Intelligent Lab. Syst., 110 (2012), pp. 1–19. (Cited on p. 430)
[150] A. Klein and P. Spreij, *On Fisher's information matrix of an ARMAX process and Sylvester's resultant matrices*, Linear Algebra Appl., 237/238 (1996), pp. 579–590. (Cited on p. 384)
[151] A. Klein and P. J. C. Spreij, *On Stein's equation, Vandermonde matrices and Fisher's information matrix of time series processes. Part* I: *The autoregressive moving average process*, Linear Algebra Appl., 329 (2001), pp. 9–47. (Cited on pp. 384, 418)
[152] A. Klein and P. Spreij, *On the solution of Stein's equation and Fisher's information matrix of an ARMAX process*, Linear Algebra Appl., 396 (2005), pp. 1–34. (Cited on pp. 384, 418)
[153] L. Knizhnerman and V. Simoncini, *Convergence analysis of the extended Krylov subspace method for the Lyapunov equation*, Numer. Math., 118 (2011), pp. 567–586. (Cited on p. 409)
[154] T. G. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500. (Cited on p. 430)
[155] N. Komaroff, *Simultaneous eigenvalue lower bounds for the Lyapunov matrix equation*, IEEE Trans. Automat. Control, 33 (1988), pp. 126–128. (Cited on p. 405)

[156] H. Kong, B. Zhou, and M.-R. Zhang, *A Stein equation approach for solutions to the Diophantine equations*, in the 2010 Chinese Control and Decision Conference (CCDC), 2010, pp. 3024–3028. (Cited on pp. 418, 428)
[157] M. Konstantinov, D. Gu, V. Mehrmann, and P. Petkov, *Perturbation Theory for Matrix Equations*, Stud. Comput. Math. 9, Elsevier, New York, 2003. (Cited on pp. 380, 388, 389, 423)
[158] D. Kressner, *Block variants of Hammarling's method for solving Lyapunov equations*, ACM Trans. Math. Software, 34 (2008), pp. 1–15. (Cited on pp. 406, 419)
[159] D. Kressner, *Memory-efficient Krylov subspace techniques for solving large-scale Lyapunov equations*, in IEEE International Symposium on Computer-Aided Control Systems, San Antonio, 2008, IEEE, pp. 613–618. (Cited on p. 401)
[160] D. Kressner, V. Mehrmann, and T. Penzl, *CTLEX: A Collection of Benchmark Examples for Continuous-Time Lyapunov Equations*, SLICOT Working Note 1999-6, www.slicot.org, 1999. (Cited on p. 429)
[161] D. Kressner, C. Schröder, and D. S. Watkins, *Implicit QR algorithms for palindromic and even eigenvalue problems*, Numer. Algorithms, 51 (2009), pp. 209–238. (Cited on p. 426)
[162] D. Kressner and C. Tobler, *Krylov subspace methods for linear systems with tensor product structure*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 1688–1714. (Cited on p. 430)
[163] D. Kressner and C. Tobler, *Low-rank tensor Krylov subspace methods for parametrized linear systems*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1288–1316. (Cited on p. 426)
[164] P. Kunkel and V. L. Mehrmann, *Differential-Algebraic Equations: Analysis and Numerical Solution*, European Mathematical Society, Zürich, Switzerland, 2006. (Cited on pp. 379, 421)
[165] P. Lancaster, *Explicit solutions of linear matrix equations*, SIAM Rev., 12 (1970), pp. 544–566. (Cited on pp. 379, 380, 387, 388)
[166] P. Lancaster, L. Lerer, and M. Tismenetsky, *Factored forms for solutions of $AX-XB=C$ and $X-AXB=C$ in companion matrices*, Linear Algebra Appl., 62 (1984), pp. 19–49. (Cited on p. 419)
[167] P. Lancaster and L. Rodman, *Algebraic Riccati equations*, Oxford University Press, Oxford, UK, 1995. (Cited on pp. 418, 430)
[168] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed., Academic Press, New York, 1985. (Cited on pp. 387, 405, 418)
[169] V. Larin and F. Aliev, *Construction of square root factor for solution of a Lyapunov matrix equation*, Systems Control Lett., 20 (1993), pp. 109–112. (Cited on p. 406)
[170] A. Laub, *Numerical linear algebra aspects of control design computations*, IEEE Trans. Automat. Control, AC-30 (1985), pp. 97–108. (Cited on p. 383)
[171] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK Users Guide: Solution of Large Scale Eigenvalue Problems by Implicitly Restarted Arnoldi Methods*, SIAM, Philadelphia, 1998. (Cited on p. 417)
[172] L. Lerer and A. C. M. Ran, *A new inertia theorem for Stein equations, inertia of invertible Hermitian block Toeplitz matrices and matrix orthogonal polynomials*, Integral Equations Oper. Theory, 47 (2003), pp. 339–360. (Cited on p. 418)
[173] J.-R. Li, *Model Reduction of Large Linear Systems via Low Rank System Gramians*, Ph.D. thesis, MIT, Cambridge, MA, 2000. (Cited on p. 407)
[174] J.-R. Li and J. White, *Low rank solution of Lyapunov equations*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 260–280. (Cited on pp. 412, 414)
[175] T. Li, P. C.-Y. Weng, E. K. Chu, and W.-W. Lin, *Large-scale Stein and Lyapunov equations, Smith method, and applications*, Numer. Algorithms, 63 (2013), pp. 727–752. (Cited on p. 420)
[176] J. Liesen and Z. Strakos, *Krylov Subspace Methods. Principles and Analysis*, Oxford University Press, Oxford, UK, 2012. (Cited on p. 405)
[177] M. M. Lin and C.-Y. Chiang, *On the T-Stein Matrix Equation $X = AX^\top B+C$*, Tech. Report 2012-10-002, National Center for Theoretical Sciences, National Tsing Hua University, Taiwan, 2012. (Cited on p. 420)
[178] Y. Lin, L. Bao, Y. Cao, and L. Zhou, *A direct method for solving projected generalized continuous-time Sylvester equations*, J. Control Theory Appl, 10 (2012), pp. 319–324. (Cited on p. 421)
[179] Y. Lin and V. Simoncini, *Minimal residual methods for large scale Lyapunov equations*, Appl. Numer. Math., 72 (2013), pp. 52–71. (Cited on pp. 397, 408, 411)
[180] Y. Lin and Y. Wei, *Condition numbers of the generalized Sylvester equation*, IEEE Trans. Automat. Control, 52 (2007), pp. 2380–2385. (Cited on p. 423)

[181] A. Lu and E. L. Wachspress, *Solution of Lyapunov equations by alternating direction implicit iteration*, Comput. Math. Appl., 21 (1991), pp. 43–58. (Cited on p. 412)
[182] D. Luenberger, *Observers for multivariable systems*, IEEE Trans. Automat. Control, 11 (1966), pp. 190–197. (Cited on pp. 383, 424)
[183] M. Mariton, *Jump Linear Systems in Automatic Control*, Marcel Dekker, New York, Basel, 1990. (Cited on p. 385)
[184] K. Meerbergen and A. Spence, *Inverse iteration for purely imaginary eigenvalues with application to the detection of Hopf bifurcations in large-scale problems*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 1982–1999. (Cited on pp. 383, 430)
[185] C. C. K. Mikkelsen, *Numerical Methods for Large Lyapunov Equations*, Ph.D. thesis, Purdue University, West Lafayette, IN, 2009. (Cited on pp. 407, 416)
[186] C. B. Moler and G. W. Stewart, *An algorithm for generalized eigenvalue problems*, SIAM J. Numer. Anal., 10 (1973), pp. 241–256. (Cited on p. 422)
[187] M. M. Monahemi, J. B. Barlow, and D. P. O'Leary, *The design of reduced order observers with precise loop transfer recovery*, AIAA J. Guidance Control Dynam., 15 (1992), pp. 1320–1326. (Cited on p. 425)
[188] C. T. Mullis and R. A. Roberts, *Synthesis of minimum roundoff noise fixed-point digital filters*, IEEE Trans. Circuits Systems, CAS-23 (1976), pp. 551–562. (Cited on p. 382)
[189] H. D. Nong, *Numerical Solutions of Matrix Equations Arising in Model Reduction of Large-Scale Linear-Time-Invariant Systems*, Ph.D. thesis, Rice University, Houston, TX, 2010. (Cited on pp. 407, 417)
[190] R. Nong and D. C. Sorensen, *A Parameter Free ADI-Like Method for the Numerical Solution of Large Scale Lyapunov Equations*, Tech. Report TR09-16, Rice University, Houston, TX, 2009. (Cited on p. 417)
[191] I. V. Oseledets and E. E. Tyrtyshnikov, *Breaking the curse of dimensionality, or how to use SVD in many dimensions*, SIAM J. Sci. Comput., 31 (2009), pp. 3744–3759. (Cited on p. 430)
[192] I. V. Oseledets and E. E. Tyrtyshnikov, *TT cross approximation for multidimensional arrays*, Linear Algebra Appl., 432 (2010), pp. 70–88. (Cited on p. 430)
[193] R. Patel, A. J. Laub, and P. Van Dooren, *Numerical Linear Algebra Techniques for Systems and Control*, IEEE Press, Piscataway, NJ, 1993. (Cited on p. 383)
[194] D. W. Peaceman and H. H. Rachford, Jr., *The numerical solution of parabolic and elliptic equations*, J. Soc. Indust. Appl. Math., 3 (1955), pp. 28–41. (Cited on p. 401)
[195] R. Peeters and P. Rapisarda, *A two-variable approach to solve the polynomial Lyapunov equation*, Systems Control Lett., 42 (2001), pp. 117–126. (Cited on p. 428)
[196] T. Penzl, *Numerical solution of generalized Lyapunov equations*, Adv. Comput. Math., 8 (1998), pp. 33–48. (Cited on pp. 421, 422)
[197] T. Penzl, *Numerische Lösung großer Lyapunov-Gleichungen*, Diss. Fakultät für Mathematik, TU Chemnitz, Logos-Verlag, Berlin, 1998. (Cited on p. 407)
[198] T. Penzl, *A cyclic low-rank Smith method for large sparse Lyapunov equations*, SIAM J. Sci. Comput., 21 (2000), pp. 1401–1418. (Cited on pp. 412, 413, 414, 420)
[199] T. Penzl, *Eigenvalue decay bounds for solutions of Lyapunov equations: The symmetric case*, Systems Control Lett., 40 (2000), pp. 139–144. (Cited on pp. 396, 405)
[200] T. Penzl, *Lyapack Users Guide*, Tech. Report SFB393/00-33, TU Chemnitz, 09107 Chemnitz, D, 2000. Available from http://www.tu-chemnitz.de/sfb393/sfb00pr.html. (Cited on pp. 412, 414, 428)
[201] V. Popov, *Hyperstability and optimality of automatic systems with several control functions*, Rev. Roum. Sci. Tech., 9 (1964), pp. 629–690. (Cited on p. 419)
[202] P. Poromaa, *Parallel algorithms for triangular Sylvester equations: Design, scheduling and scalability issues*, in Applied Parallel Computing Large Scale Scientific and Industrial Problems, Lecture Notes in Comput. Sci. 1541, Springer, 1998, pp. 438–446. (Cited on p. 429)
[203] C. E. Powell and D. J. Silvester, *Preconditioning steady-state Navier–Stokes equations with random data*, SIAM J. Sci. Comput., 34 (2012), pp. A2482–A2506. (Cited on p. 386)
[204] H. M. Power, *A note on the matrix equation $A'LA - L = -K$*, IEEE Trans. Automat. Control, 14 (1969), pp. 411–412. (Cited on p. 419)
[205] E. S. Quintana-Ortí and R. van de Geijn, *Specialized parallel algorithms for solving Lyapunov and Stein equations*, J. Parallel Distrib. Comput., 61 (2001), pp. 1489–1504. (Cited on p. 429)
[206] E. S. Quintana-Ortí and R. van de Geijn, *Formal derivation of algorithms: The triangular Sylvester equation*, ACM Trans. Math. Software, 29 (2003), pp. 218–243. (Cited on p. 429)

[207] A. C. M. Ran, M. C. B. Reurings, and L. Rodman, *A perturbation analysis for nonlinear selfadjoint operator equations*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 89–104. (Cited on p. 430)

[208] H. Ren and R. Dekany, *Fast wave-front reconstruction by solving the Sylvester equation with the alternating direction implicit method*, Optics Express, 12 (2004), pp. 3279–3296. (Cited on p. 384)

[209] M. Robbé and M. Sadkane, *A convergence analysis of GMRES and FOM methods for Sylvester equations*, Numer. Algorithms, 30 (2002), pp. 71–89. (Cited on pp. 393, 394)

[210] J. D. Roberts, *Linear model reduction and solution of the algebraic Riccati equation by use of the sign function*, Internat. J. Control, 32 (1980), pp. 677–687. (Cited on p. 407)

[211] W. E. Roth, *The equations $AX - YB = C$ and $AX - XB = C$ in matrices*, Proc. Amer. Math. Soc., 3 (1952), pp. 392–396. (Cited on pp. 387, 422, 423)

[212] A. Ruhe, *Rational Krylov sequence methods for eigenvalue computation*, Linear Algebra Appl., 58 (1984), pp. 391–405. (Cited on p. 400)

[213] Y. Saad, *Numerical solution of large Lyapunov equations*, in Signal Processing, Scattering, Operator Theory, and Numerical Methods. Proceedings of the International Symposium MTNS-89, Vol. III, M. A. Kaashoek, J. H. van Schuppen, and A. C. Ran, eds., Birkhäuser, Boston, 1990, pp. 503–511. (Cited on pp. 398, 408)

[214] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003. (Cited on pp. 380, 392, 394, 401)

[215] J. Sabino, *Solution of Large-Scale Lyapunov Equations via the Block Modified Smith Method*, Ph.D. thesis, Rice University, Houston, TX, 2006. (Cited on pp. 395, 396, 406, 407)

[216] M. Sadkane, *A low-rank Krylov squared Smith method for large-scale discrete-time Lyapunov equations*, Linear Algebra Appl., 436 (2012), pp. 2807–2827. (Cited on p. 420)

[217] B. Schäfer-Bung, C. Hartmann, B. Schmidt, and C. Schütte, *Dimension reduction by balanced truncation: Application to light-induced control of open quantum systems*, J. Chem. Phys., 135 (2011), 014112. (Cited on p. 385)

[218] W. H. A. Schilders, H. A. van der Vorst, and J. Rommes, *Model Order Reduction: Theory, Research Aspects and Applications*, Springer-Verlag, Berlin, Heidelberg, 2008. (Cited on pp. 381, 383)

[219] P. J. Schmid, *Nonmodal stability theory*, Annu. Rev. Fluid Mech., 39 (2007), pp. 129–162. (Cited on p. 384)

[220] S. D. Shank and V. Simoncini, *Krylov subspace methods for large-scale constrained Sylvester equations*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1448–1463. (Cited on p. 425)

[221] F. C. Silva and R. Simões, *On the Lyapunov and Stein equations*, I, Linear Algebra Appl., 420 (2007), pp. 329–338. (Cited on pp. 405, 419)

[222] F. C. Silva and R. Simões, *On the Lyapunov and Stein equations*, II, Linear Algebra Appl., 426 (2007), pp. 305–311. (Cited on pp. 405, 419)

[223] V. Sima, *Algorithms for Linear-Quadratic Optimization*, Marcel Dekker, New York, 1996. (Cited on pp. 390, 406, 428)

[224] V. Sima and P. Benner, *Experimental evaluation of new SLICOT solvers for linear matrix equations based on the matrix sign function*, in Proceedings of the 9th IEEE International Conference on Computer-Aided Control Systems (CACSD 2008), 2008, pp. 601–606. (Cited on p. 416)

[225] V. Simoncini, *On the numerical solution of $AX - XB = C$*, BIT, 36 (1996), pp. 814–830. (Cited on pp. 388, 393, 397, 400)

[226] V. Simoncini, *A new iterative method for solving large-scale Lyapunov matrix equations*, SIAM J. Sci. Comput., 29 (2007), pp. 1268–1288. (Cited on pp. 409, 414)

[227] V. Simoncini, *The extended Krylov subspace for parameter dependent systems*, Appl. Numer. Math., 60 (2010), pp. 550–560. (Cited on p. 395)

[228] V. Simoncini and V. Druskin, *Convergence analysis of projection methods for the numerical solution of large Lyapunov equations*, SIAM J. Numer. Anal., 47 (2009), pp. 828–843. (Cited on p. 408)

[229] V. Simoncini and M. Sadkane, *Arnoldi-Riccati method for large eigenvalue problems*, BIT, 36 (1996), pp. 579–594. (Cited on p. 383)

[230] V. Simoncini and D. B. Szyld, *Recent computational developments in Krylov subspace methods for linear systems*, Numer. Linear Algebra Appl., 14 (2007), pp. 1–59. (Cited on pp. 392, 415)

[231] J. R. Singler, *Approximate low rank solutions of Lyapunov equations via proper orthogonal decomposition*, in 2008 American Control Conference, Seattle, WA, 2008, pp. 267–272. (Cited on p. 418)

[232] M. Slowik, P. Benner, and V. Sima, *Evaluation of the linear matrix equation solvers in SLICOT*, JNAIAM J. Numer. Anal. Ind. Appl. Math., 2 (2007), pp. 11–34. (Cited on pp. 390, 428)

[233] R. A. Smith, *Matrix equation* $XA+BX=C$, SIAM J. Appl. Math., 16 (1968), pp. 198–201. (Cited on pp. 401, 419)

[234] D. C. Sorensen and A. C. Antoulas, *The Sylvester equation and approximate balanced reduction*, Linear Algebra Appl., 351–352 (2002), pp. 671–700. (Cited on pp. 382, 392)

[235] D. Sorensen and Y. Zhou, *Bounds on Eigenvalue Decay Rates and Sensitivity of Solutions to Lyapunov Equations*, Tech. Report 02-07, Rice University, Houston, TX, 2002. (Cited on p. 405)

[236] D. C. Sorensen and Y. Zhou, *Direct methods for matrix Sylvester and Lyapunov equations*, J. Appl. Math., 2003 (2003), pp. 277–303. (Cited on p. 390)

[237] G. Starke and W. Niethammer, *SOR for* $AX-XB=C$, Linear Algebra Appl., 154–156 (1991), pp. 355–375. (Cited on p. 403)

[238] G. W. Stewart and J-G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990. (Cited on pp. 383, 389, 421, 423, 424)

[239] T. Stykel, *Numerical solution and perturbation theory for generalized Lyapunov equations*, Linear Algebra Appl., 349 (2002), pp. 155–185. (Cited on p. 421)

[240] J. J. Sylvester, *Sur l'equations en matrices* $px=xq$, C.R. Acad. Sci. Paris, 99 (1884), pp. 67–71,115–116. (Cited on p. 378)

[241] V. L. Syrmos, *Disturbance decoupling using constrained Sylvester equations*, IEEE Trans. Automat. Control, 39 (1994), pp. 797–803. (Cited on p. 425)

[242] A. Talbot, *The evaluation of integrals of products of linear system responses, Part* I, Quart. J. Mech. Appl. Math., 12 (1959), pp. 488–503. (Cited on p. 428)

[243] N. Truhar, Z. Tomljanović, and R.-C. Li, *Analysis of the solution of the Sylvester equation using low-rank ADI with exact shifts*, Systems Control Lett., 59 (2010), pp. 248–257. (Cited on p. 389)

[244] N. Truhar and K. Veselić, *Bounds on the trace of a solution to the Lyapunov equation with a general stable matrix*, Systems Control Lett., 56 (2007), pp. 493–503. (Cited on p. 413)

[245] C.-C. Tsui, *A new approach to robust observer design*, Internat. J. Control, 47 (1988), pp. 745–751. (Cited on p. 425)

[246] A. van den Boom, A. Brown, F. Dumortier, A. Geurts, S. Hammarling, R. Kool, M. Vanbegin, P. Van Dooren, and S. Van Huffel, *SLICOT, a subroutine library for control and system theory*, in Preprints IFAC Symp. on CADCS, Swansea, UK, 1991, pp. 89–94. (Cited on pp. 390, 428)

[247] P. Van Dooren, *Structured linear algebra problems in digital signal processing*, in Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms, NATO Series F, Springer, New York, 1991, pp. 361–384. (Cited on p. 381)

[248] P. Van Dooren, *Reduced order observers: A new algorithm and proof*, System Control Lett., 4 (1984), pp. 243–251. (Cited on pp. 424, 425)

[249] P. Van Dooren and M. Verhaegen, *Condensed forms for efficient time-invariant Kalman filtering*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 516–530. (Cited on p. 425)

[250] B. Vandereycken and S. Vandewalle, *A Riemannian optimization approach for computing low-rank solutions of Lyapunov equations*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2553–2579. (Cited on p. 418)

[251] A. Varga, *A note on Hammarling's algorithm for the discrete Lyapunov equation*, Systems Control Lett., 15 (1990), pp. 273–275. (Cited on p. 419)

[252] D. Vasilyev and J. White, *A more reliable reduction algorithm for behavior model extraction*, in Proceedings of the IEEE/ACM International Conference on Computer-Aided Design, 2005, pp. 813–820. (Cited on p. 417)

[253] Yu. O. Vorontsov and Kh. D. Ikramov, *A numerical algorithm for solving the matrix equation* $AX+X^TB=C$, Comput. Math. Math. Phys., 51 (2011), pp. 691–698. (Cited on p. 426)

[254] E. L. Wachspress, *Extended application of alternating direction implicit iteration model problem theory*, J. Soc. Indust. Appl. Math., 11 (1963), pp. 994–1016. (Cited on pp. 384, 391)

[255] E. L. Wachspress, *Iterative solution of the Lyapunov matrix equation*, Appl. Math. Lett., 1 (1988), pp. 87–90. (Cited on pp. 402, 412)

[256] F. Wan, *An in-core finite difference method for separable boundary value problems on a rectangle*, Stud. Appl. Math., 52 (1973), pp. 103–113. (Cited on p. 391)

[257] Q.-W. Wang, J.-H. Sun, and S.-Z. Li, *Consistency for bi(skew)symmetric solutions to systems of generalized Sylvester equations over a finite central algebra*, Linear Algebra Appl., 353 (2002), pp. 169–182. (Cited on p. 424)

[258] D. S. Watkins, *The Matrix Eigenvalue Problem. GR and Krylov Subspace Methods*, SIAM, Philadelphia, 2007. (Cited on p. 391)

[259] H. K. Wimmer, *The matrix equation $X - AXB = C$ and an analogue of Roth's theorem*, Linear Algebra Appl., 109 (1988), pp. 145–147. (Cited on p. 419)

[260] H. K. Wimmer and A. D. Ziebur, *Remarks on inertia theorems for matrices*, Czechoslovak Math. J., 25 (1975), pp. 556–561. (Cited on p. 418)

[261] H. K. Wimmer, *Linear matrix equations, controllability and observability, and the rank of solutions*, SIAM J. Matrix Anal., 9 (1988), pp. 570–578. (Cited on p. 388)

[262] H. K. Wimmer, *Explicit solutions of the matrix equation $\sum A^i X D_i = C$*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1123–1130. (Cited on p. 388)

[263] H. K. Wimmer, *The generalized Sylvester equation in polynomial matrices*, IEEE Trans. Automat. Control, 41 (1996), pp. 1372–1376. (Cited on p. 422)

[264] A.-G. Wu, G.-R. Duan, and Y. Xue, *Kronecker maps and Sylvester-polynomial matrix equations*, IEEE Trans. Automat. Control, 52 (2007), pp. 905–910. (Cited on p. 425)

[265] A.-G. Wu, G. Feng, G.-R. Duan, and W.-J. Wu, *Closed-form solutions to Sylvester-conjugate matrix equations*, Comput. Math. Appl., 60 (2010), pp. 95–111. (Cited on p. 427)

[266] A.-G. Wu, J. Hu, and G.-R. Duan, *Solutions to the matrix equation $AX - EXF = BY$*, Comput. Math. Appl., 58 (2009), pp. 1891–1900. (Cited on p. 423)

[267] A.-G. Wu, Y. Sun, and G. Feng, *Closed-form solution to the non-homogeneous generalised Sylvester matrix equation*, IET Control Theory Appl., 4 (2010), pp. 1914–1921. (Cited on p. 423)

[268] B. Zhou and G.-R. Duan, *Parametric solutions to the generalized discrete Sylvester matrix equation $MXN - X = TY$ and their applications*, IMA J. Math. Control Inform., 26 (2009), pp. 59–78. (Cited on p. 423)

[269] B. Zhou, G.-R. Duan, and Z.-Y. Li, *Gradient based iterative algorithm for solving coupled matrix equations*, Systems Control Lett., 58 (2009), pp. 327–333. (Cited on pp. 418, 423, 424)

[270] B. Zhou, G.-R. Duan, and Z.-Y. Li, *A Stein matrix equation approach for computing coprime matrix fraction description*, IET Control Theory Appl., 3 (2009), pp. 691–700. (Cited on p. 383)

[271] B. Zhou, J. Lam, and G.-R. Duan, *Toward Solution of Matrix Equation $X = Af(X)B + C$*, Tech. Report, arXiv:1211.0346, 2012. (Cited on p. 420)

[272] Y. Zhou, *Numerical Methods for Large Scale Matrix Equations with Applications in LTI System Model Reduction*, Ph.D. thesis, Rice University, Houston, Texas, 2002. (Cited on pp. 406, 407)