Numerical solution of singular Lyapunov equations

Eric King-wah Chu, Daniel B. Szyld and Jieyong Zhou

Report 16-06-02
June 2016. Revised September 2019 and April 2020.

Department of Mathematics
Temple University
Philadelphia, PA 19122

This report is available in the World Wide Web at
http://www.math.temple.edu/~szyld

NUMERICAL SOLUTION OF SINGULAR LYAPUNOV EQUATIONS*

ERIC KING-WAH CHU†, DANIEL B. SZYLD‡, AND JIEYONG ZHOU§

Abstract. We consider the numerical solution of large scale singular (continuous time) Lyapunov equations of the form $AX + XA^\top + BB^\top = 0$, where A is semi-stable, i.e., its spectrum is contained in the left half plane, with the exception of a few semi-simple eigenvalues at zero. We also consider the case of a few semi-simple eigenvalues on the imaginary axis. Given the invariant subspaces corresponding to those few semi-simple eigenvalues, our approach is to find a low-rank approximation in the complement of those invariant subspaces. We thus obtain a low-rank approximation to the least squares solution. For large systems, we propose an appropriate space to be used with projection methods. Selected numerical examples are provided illustrating our proposed approach.

Key words. Lyapunov equation, projection method, singular equation least squares solution, Krylov subspace, Sylvester equation, Stein equation

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

1. Introduction.

We are interested in the solution of the singular (continuous time) Lyapunov equation (sLE):

$$AX + XA^\top + BB^\top = 0, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$ is large, sparse and semi-stable, and $B \in \mathbb{R}^{n\times p}$ is full rank with $p \ll n$. A matrix A is called stable if its eigenvalues, denoted by $\lambda(A)$, lie on the open left half complex plane $\mathbb{C}_-$, i.e., if $Re(\lambda(A)) < 0$; it is called Lyapunov-stable if its eigenvalues lie in the closed half complex plane $\mathbb{C}_- \cup i\mathbb{R}$, and it is called semi-stable if most of its eigenvalues lie on $\mathbb{C}_-$, except for a few semi-simple eigenvalues (also called non-degenerate eigenvalues) at zero. We are interested in a low-rank approximation to the solution X of (1.1); in other words, we are seeking $Z \in \mathbb{R}^{n\times q}$ such that $X \approx ZZ^\top$ with its rank $q \ll n$.

It is well known that the Lyapunov operator $L(\cdot) = A(\cdot) + (\cdot)A^\top$ is nonsingular if and only if A is stable. We shall consider the more general situation when either A is Lyapunov stable or semi-stable, making L singular without any $\lambda(A)$ crossing over into the open right half plane. We are mainly concerned with the case of when A is semi-stable, although most results carry through to the Lyapunov-stable case. We do not treat explicitly the case of degenerate eigenvalues but our results can be used in this case as well after appropriate modifications.

Semi-stable matrices play an important role when analyzing autonomous systems [9], matrix second-order systems [4] which is related to modeling vibration [32], and certain nonlinear systems [5, 6]. In the last few years, a theory for optimal semi-stable control theory for linear and nonlinear dynamical systems was developed [24, 28, 41], and applied to several systems. These include systems with a continuum of equilibria and have many applications in mechanical systems having rigid body modes, chemical reaction systems [21], compartmental systems [22, 23], isospectral matrix dynamical systems and dynamical network systems [30, 31, 46]. They cover a broad spectrum of applications including cooperative control of unmanned air vehicles, autonomous

*This version dated 22 April 2020

†School of Mathematics, Monash University, 9 Rainforest Walk, Melbourne, Victoria 3800, Australia; eric.chu@monash.edu

‡Department of Mathematics, Temple University, 1805 N Broad Street, Philadelphia, PA 19122 (szyld@temple.edu). Supported in part by the U.S. National Science Foundation under grant DMS-1418882 and the U.S. Department of Energy under grant DE-SC 0016578.

§Mathematics School, Shanghai Key Laboratory of Financial Information Technology, Shanghai University of Finance and Economics, Shanghai, 200433, China (jieyongzhou@gmail.com). Supported in part by the Shanghai Municipal Education Commission.

underwater vehicles, distributed sensor networks, air and ground transportation systems, swarms of air and space vehicle formations and congestion control in communication networks.

When A is stable (thus the Lyapunov operator is nonsingular) and n is small, the well-known direct method by Bartels and Stewart [1] for Sylvester equations can be applied to (1.1). For large and sparse matrices, Saad proposed a Galerkin projection method [48, 49] onto the block Krylov subspace $\mathcal{K}_k(A,B) = \mathcal{R}([B, AB, A^2B, \ldots, A^{k-1}B])$ (where $\mathcal{R}(\cdot)$ denotes the range). See, e.g., [34, 38], for other block Krylov projections for Lyapunov equations, [14] for Sylvester equations, and [52] for a more recent survey of methods for matrix equations. For a general treatment of block Krylov subspaces, see the survey [20], and references therein.[1] See also the earlier paper by Hu and Reichel [27], where a projection with a standard Krylov method is used. Another approach is to use the ADI method [3, 47].

In [38], it is proposed that the projection method be used with an extended Krylov subspace of the form $\mathcal{K}_k(A,B)+\mathcal{K}_k(A^{-1},A^{-1}B)$, and it is shown numerically that this approach produces approximations to the solutions of (1.1) efficiently. In [13], the rational Krylov subspace (RKS) was employed in the projection method; see also [12]. While the projection with extended Krylov is often cheaper, the use of the RKS is superior to the extended Krylov subspace when, for instance, the field of values of A is very close to the imaginary axis [51]. In most nonsingular cases, projection with an extended or a rational Krylov subspace is expected to be effective.

Following [49], the Galerkin projection for any of these spaces consists of computing a matrix $V_k \in \mathbb{R}^{n\times\nu_k}$ whose ν_k columns form an orthonormal basis of the appropriate Krylov subspace. Then, the approximate solution is of the form $X_k = V_k Y_k V_k^T$ with $Y_k \in \mathbb{R}^{\nu_k\times\nu_k}$ and the corresponding residual is

$$R_k = AX_k + X_kA^\top + BB^\top. \tag{1.2}$$

The Galerkin condition that the residual be orthogonal to all elements of the form $V_kY_kV_k^\top$ is equivalent to the lower order Lyapunov equation

$$A_kY_k + Y_kA_k^\top + B_kB_k^\top = 0, \tag{1.3}$$

with $A_k \equiv V_k^\top AV_k$ and $B_k \equiv V_k^\top B$, which can be solved with the Bartels-Stewart method [1] mentioned above. Here orthogonality is with respect to the Frobenius inner product. In [42] a minimal residual solution in the space $\mathcal{R}(V_k)$ is explored.

A more detailed algorithm with this projection approach is given as Algorithm 5.2, where (1.3) is solved in the minimal residual sense, since in our setting, the lower order Lyapunov equation (1.3) is in general singular.

In the case of a semi-stable A, (symmetric) solutions to the Lyapunov equation (1.1) exist under certain conditions, and the least squares solution is related to the $\mathcal{H}_2$ optimal solution; see [24, 29].

When considering numerical methods for the singular case, there are very few options. One could consider projection with RKS, although this approach was designed for a nonsingular equation, and since the singular system may produce ill-conditioning intermediate linear systems, this may not work as well; see also Remark 3.1 and Section 7 for an illustrative numerical experiment. The application of the method of [42] to the sLE (1.1) does not seem to be direct; see further Remark 3.3.

What we propose here for the numerical solution of the Lyapunov equation (1.1) when A is semi-stable (or Lyapunov-stable) is to isolate the singularity of the Lyapunov operator, corresponding to the few eigenvalues of A at zero (or on the imaginary axis), and to use Krylov projection with a particular space for the rest of the operator. Alternatively, we can use Krylov

[1] We also mention the recent paper [16], where a general view of block Krylov methods is presented.

projection directly on the original equation (1.1) using an appropriate space. Part of our contribution is to define such a subspace, and show how it can be computed and used in this context.

To fix ideas let $\ell = \dim(\mathcal{N}(A))$, the dimension of the null space of A, which is relatively small, i.e., $\ell \ll n$. Let $\mathcal{N}(A) = \mathcal{R}(P_1)$ with $P_1 \in \mathbb{R}^{n\times\ell}$ and $P = [P_1, P_2]$ being orthogonal. Thus, at least in theory, the space can be partitioned into two parts, and we can decouple the solution X into the corresponding four parts. With $X_{ij} = P_i X P_j^\top$ $(i, j = 1, 2)$, the solution to (1.1) can then be written as

$$X = P_1 X_{11} P_1^\top + P_1 X_{12} P_2^\top + P_2 X_{12}^\top P_1^\top + P_2 X_{22} P_2^\top, \tag{1.4}$$

where X_{11} and X_{22} are symmetric. As $\ell \ll n$, the parts on the right-hand-side of (1.4) involving X_{11} and X_{12} are low-rank. We shall show later that X_{22} satisfies a Lyapunov equation defined by a stable matrix with a low-rank constant term. This implies that, although X_{22} is usually full rank, it can be approximated by a low-rank matrix. Thus, the overall solution X can be approximated by a low-rank matrix $ZZ^\top$ with $Z \in \mathbb{R}^{n\times q}$ $(q \ll n)$ and we shall identify the appropriate (Krylov) subspace for the corresponding projection method.

The paper is organized as follows. In Section 2, we consider the transformation of the Lyapunov equation (1.1) using a Schur-like decomposition of A. The Lyapunov equation (1.1) becomes a small singular Lyapunov equation in X_{11}, a uniquely solvable Sylvester equation in X_{12} and a uniquely solvable Lyapunov equation in X_{22}. Solution of the Lyapunov equation (1.1) via the transformed equations is then considered. This works well for small size matrices. For a larger n, we consider Galerkin projection to an appropriate space (Section 3). In Section 4 we make some comments on the stability of the matrices in the projected equations, while in Section 5 we detail the algorithms for large matrices. We make some remarks on the solution of the singular Stein equation $-X + AXA^\top + BB^\top = 0$ in Section 6. Some illustrative numerical examples are presented in Section 7, and we conclude in Section 8.

Our contribution consists of treating the solution of a singular Lyapunov systems as the sum of different parts, each of which can be computed by solving different matrix equations. Each part is either low-rank, or can be approximated by a low-rank matrix. We propose an appropriate space to be used with a projection method. While these two ideas do not solve the singular systems in all situations, we show how these proposals can be used successfully in many practical cases. In those cases, we provide both computational details and illustrative numerical examples.

For notation, $\|\cdot\|$ indicates the 2-norm, while $\|\cdot\|_F$ corresponds to the Frobenius norm, and $\Lambda(\cdot)$ denotes the spectrum.

2. Schur-Like Decomposition of A. Let us consider the slightly more general case of a Lyapunov stable matrix A with $\ell \ll n$ semi-simple eigenvalues on the imaginary axis. Thus, we can write

$$A = P \begin{bmatrix} \Omega & A_{12} \\ 0 & A_{22} \end{bmatrix} P^\top, \tag{2.1}$$

where $P = [P_1, P_2]$ is unitary, $A_{22} = P_2^\top A P_2$ is stable (with all eigenvalues in $\mathbb{C}_-$), and thus invertible, and $\Omega \in \mathbb{R}^{\ell\times\ell}$ $(\ell \ll n)$ has only semi-simple eigenvalues on the imaginary axis. Note that the decomposition in (2.1) for a given A is totally determined by P_1. When Ω and A_{22} are in Schur form, which is not essential in the paper, then (2.1) is the Schur decomposition of A.

If $\Omega = 0$ (the semi-stable case), then $\mathcal{N}(A) = \mathcal{R}(P_1)$, which is relatively small in dimension, and thus, it can be computed efficiently; see, e.g., [19]. After P_1 is computed (possibly by inverse simultaneous iteration), P_2 can then be obtained via the QR decomposition of P_1 [18]. We note that neither P_1, P_2, nor A_{22} needs to be computed explicitly. The matrix-vector product with P_i, or its transpose, can be performed by using the stored angles for the Givens rotations, or

the parameters of the Householder reflections, and the product with A_{22} using its product form $P_2^\top A P_2$.

When $\Omega \neq 0$, P_1 is more difficult to compute. One possibility is to apply Newton's method to $\left[(A^2+\rho I)x, z^\top x - 1\right] = 0$ for some nonzero $z \in \mathbb{R}^n$ used for the scaling of x, to obtain the eigen-pair (ρ, x) with $\rho > 0$ and $x \neq 0$. Another approach is to apply Arnoldi's method on A coupled with some refinement technique. For the purpose of this paper, we shall assume that P_1 is readily available; and we shall see in Section 7 that in many cases this assumption can be satisfied. For the general case, the accurate estimation of P_1 is a nontrivial task in eigenvalue problems and this topic is outside the scope of this paper.

With $P = [P_1, P_2]$, the sLE (1.1) is transformed into three components as shown in our main theorem below.

THEOREM 2.1 (Singular Lyapunov Equations). *Given $P = [P_1, P_2]$, one can write A as in* (2.1)*. Then, the sLE* (1.1) *is equivalent to the following three equations*

$$\Omega X_{11} + X_{11}\Omega^\top + A_{12}X_{12}^\top + X_{12}A_{12}^\top + B_1 B_1^\top = 0, \tag{2.2a}$$

$$\Omega X_{12} + X_{12}A_{22}^\top + A_{12}X_{22} + B_1 B_2^\top = 0, \tag{2.2b}$$

$$A_{22}X_{22} + X_{22}A_{22}^\top + B_2 B_2^\top = 0, \tag{2.2c}$$

where $X_{ij} \equiv P_i^\top X P_j$ and $B_i \equiv P_i^\top B$ $(i,j = 1,2)$. Consequently, with the symmetric positive semidefinite X_{22}, and with X_{12} uniquely determined by the Lyapunov equation (2.2c) *and the Sylvester equation* (2.2b)*, respectively, we have that if X_{11} is the least squares solution to* (2.2a)*, then $X = \sum_{i,j} P_i X_{ij} P_j^\top$ is the least squares solution to* (1.1) *in Frobenius norm. Furthermore, the residuals of* (1.1) *and* (2.2a) *are identical in norm.*

Proof. Pre- and post-multiply the Lyapunov equation (1.1) by $P^\top$ and P respectively, and denoting $\widetilde{X} \equiv P^\top X P = \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^\top & X_{22} \end{bmatrix}$ where X_{11} and X_{22} are symmetric and $P^\top B = [B_1^\top, B_2^\top]^\top$, we obtain the equivalent equation

$$\begin{bmatrix} \Omega & A_{12} \\ 0 & A_{22} \end{bmatrix} \widetilde{X} + \widetilde{X} \begin{bmatrix} \Omega^\top & 0 \\ A_{12}^\top & A_{22}^\top \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} [B_1^\top, B_2^\top] = 0,$$

which is equivalent to (2.2).

Equation (2.2c) is a Lyapunov equation with A_{22} being stable, possessing a unique symmetric positive semidefinite solution X_{22}. In turn, (2.2b) is a uniquely solvable Sylvester equation in X_{12} as $\Lambda(\Omega) \cap \Lambda(-A_{22}) = \emptyset$. This leaves the small sLE (2.2a) which is solved in the least squares sense for X_{11}. As (2.2b) and (2.2c) are consistent, the corresponding residuals of (1.1) and (2.2a) are identical in norm. From $\|X\|_F^2 = \sum_{i,j} \|X_{ij}\|_F^2$ with X_{22}, X_{12} and $X_{21} = X_{12}^\top$ uniquely determined and X_{11} being the least squares solution of (2.2a), the solution $X = \sum_{i,j} P_i X_{ij} P_j^\top$ is the least squares solution to (1.1) in Frobenius norm. □

Thus, in principle, we have a road map for the solution of the singular Lyapunov equation (1.1). We obtain the least squares solution by solving first the (nonsingular) Lyapunov equation (2.2c). We then solve the Sylvester equation (2.2b). Lastly, we solve the small singular Lyapunov equation (2.2a) in a direct manner.

REMARK 2.1 (Special cases). When $\Omega = 0$, Theorem 2.1 states that the residual of (1.1) for the least squares solution equals that of (2.2a), of magnitude $\|A_{12}X_{12}^\top + X_{12}A_{12}^\top + B_1 B_1^\top\|$.

When A is symmetric, A_{22} and Ω are symmetric and $A_{12} = 0$. This forces $\Omega = 0$ and (2.2b), (2.2c) are reduced to

$$X_{12}A_{22} + B_1 B_2^\top = 0, \quad A_{22}X_{22} + X_{22}A_{22} + B_2 B_2^\top = 0.$$

The first equation (2.2a) becomes the inconsistent equation $B_1 B_1^\top = 0$. The residual for the least squares solution is thus $B_1 B_1^\top$. As X_{11} can in principle be arbitrary, and thus the minimum norm solutions is $X_{11} = 0$.

2.1. Algorithm for small sLEs. We have the following algorithm for small sLEs, based on the equivalent formulation (2.2), with the dominant flop counts summarized at the end of each step, where full matrices are assumed. For the Lyapunov equation (2.2c) and the Sylvester equation (2.2b). the Bartels-Stewart Algorithm [1], [18, Page 398] is used. Kronecker product can be applied to (2.2a) before the resulting linear system is solved in the least squares sense by the Householder QR algorithm [18, p. 293]. Using other methods or the SVD require $O(\ell^6)$ flops [18], which is negligible when $\ell \ll n$.

ALGORITHM 2.1.

1. For a given $P = [P_1, P_2]$, let $\Omega = P_1^\top A P_1 \in \mathbb{R}^{\ell\times\ell}$, $A_{12} = P_1^\top A P_2 \in \mathbb{R}^{\ell\times(n-\ell)}$, $A_{22} = P_2^\top A P_2 \in \mathbb{R}^{(n-\ell)\times(n-\ell)}$, $B_1 = P_1^\top B \in \mathbb{R}^{\ell\times p}$ and $B_2 = P_2^\top B \in \mathbb{R}^{(n-\ell)\times p}$; $2n^3$ flops.
2. Solve the Lyapunov equation (2.2c) for X_{22}; $2(n-\ell)^3$ flops.
3. Solve the Sylvester equation (2.2b) for X_{12}; $\ell(n-\ell)n$ flops.
4. Solve (2.2a) in the least squares sense for X_{11}; $O(\ell^6)$ flops.
5. Retrieve the solution $X = P \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^\top & X_{22} \end{bmatrix} P^\top$; $3n^3$ flops.

The total flop count for Algorithm 2.1 is $7n^3$ flops, ignoring $O(\ell^6)$ or other constant terms (while assuming $\ell \ll n$) and any saving from the Lyapunov equations (2.2c) and (2.2a) (which are defined respectively by one matrix operator instead of two for a general Sylvester equation, as in (2.2b)), or the symmetry of X_{11} and X_{22}.

2.2. Inconsistency. We would like make some comments on the generic case where (1.1) is inconsistent. We consider several cases. When $\Omega = 0$, X_{11} becomes free and (2.2a) is inconsistent, and the least norm solution $X_{11} = 0$ produces the residual of $A_{12}X_{12}^\top + X_{12}A_{12}^\top + B_1 B_1^\top$.

For illustration, consider the simple case when $\Omega = \begin{bmatrix} 0 & \beta \\ -\beta & 0 \end{bmatrix} \neq 0$, and (2.2a) looks like

$$\begin{bmatrix} 0 & \beta \\ -\beta & 0 \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} + \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix}\begin{bmatrix} 0 & -\beta \\ \beta & 0 \end{bmatrix} + \begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix} = 0$$

$$\Leftrightarrow \begin{bmatrix} 2\beta & 0 \\ 2\beta & 0 \\ 0 & \beta \end{bmatrix}\begin{bmatrix} x_{12} \\ x_{22} - x_{11} \end{bmatrix} + \begin{bmatrix} q_{11} \\ -q_{22} \\ q_{12} \end{bmatrix} = 0.$$

This clearly shows the generic inconsistency of (1.1), except when $q_{11} = -q_{22}$, as well as the extra freedom in $x_{22} - x_{11}$.

Rewrite (2.2a) as $\Omega X_{11} + X_{11}\Omega^\top = R \equiv -(A_{12}X_{12}^\top + X_{12}A_{12}^\top + B_1 B_1^\top)$. Using the Kronecker product $\otimes$, this equation is equivalent to $(I \otimes \Omega + \Omega \otimes I)v(X_{11}) = v(R)$, where $v(M)$ stacks the columns of the matrix M. The least squares symmetric solution X_{11} can then be obtained by the standard techniques, such as the QR or the SVD decompositions [18], when ℓ is small.

We note that when A is symmetric, then we are in the same situation described in Remark 2.1.

3. Choice of the Krylov Subspace. As we discussed in the previous section, for small to medium size n, (2.2) leads to the solution of (1.1). For large scale problems, we shall identify the appropriate subspace to be used in conjunction with a projection method.

From (1.4), (2.2b) and (2.2c), it follows that X is spanned by the columns of $P_1 \in \mathbb{R}^{n\times\ell}$, $P_2 X_{22} \in \mathbb{R}^{n\times(n-\ell)}$ and $P_2 X_{12}^\top \in \mathbb{R}^{n\times\ell}$. This indicates that any space where we would look for the solution must include $\mathcal{R}(P_1)$. Next, we observe that since X_{22} is the solution of the Lyapunov

equation (2.2c), we can thus consider approximating it in the Krylov subspace $\mathcal{K}_k(A_{22}, B_2) = \mathcal{K}_k(P_2^\top A P_2, P_2^\top B)$. Therefore $P_2 X_{22}$ can be approximated by elements in $P_2\mathcal{K}_k(A_{22}, B_2) = \mathcal{K}_k(P_2P_2^\top A, P_2P_2^\top B)$. Since X_{12} is the solution of the Lyapunov equation (2.2b), it is also sensible to approximate X_{12} by the same space as X_{22}, namely $\mathcal{K}_k(A_{22}, B_2)$; see, e.g., [52].

Let V_0 be the $n \times 2\ell$ matrix with orthonormal columns spanning $\mathcal{R}([P_1, B])$, e.g., obtained computing the QR factorization of $[P_1, B]$. As a consequence of the discussion in the previous paragraph, and using the fact that $P_1P_1^\top + P_2P_2^\top = I$, we aim to approximate the overall solution X as in (1.4) in the following space.

$$\begin{aligned}
\mathcal{S}_k &= \mathcal{R}(P_1) + \mathcal{K}_k(P_2P_2^\top A, P_2P_2^\top B) \\
&= \mathcal{R}(P_1) + \mathcal{R}([P_2P_2^\top B, (P_2P_2^\top A)P_2P_2^\top B, \cdots, (P_2P_2^\top A)^{m-1}P_2P_2^\top B]) \\
&= \mathcal{R}(P_1) + P_2P_2^\top \cdot \mathcal{R}([B, (AP_2P_2^\top)B, \cdots, (AP_2P_2^\top)^{m-1}B]) \\
&= \mathcal{R}(P_1) + P_2P_2^\top \cdot \mathcal{K}_k(AP_2P_2^\top, B) \\
&= \mathcal{R}(P_1) + (P_1P_1^\top + P_2P_2^\top) \cdot \mathcal{K}_k(AP_2P_2^\top, B) \\
&= \mathcal{R}(P_1) + \mathcal{K}_k(AP_2P_2^\top, B) = \mathcal{K}_k(AP_2P_2^\top, V_0),
\end{aligned} \tag{3.1}$$

where the third to last equality follows from the fact that $P_1P_1^\top \cdot \mathcal{K}_k(AP_2P_2^\top, B) \subseteq \mathcal{R}(P_1)$.

Some comments on this space $\mathcal{S}_k$ are in order. We may compute its basis using any of the representations in (3.1), depending on the particular structures in our sLE. Also, as already mentioned, we do not need to explicitly compute P_2, in fact, we can use the identity $P_2P_2^\top = I - P_1P_1^\top$ and compute the products with P_1. Lastly, we point out that spaces such as (3.1) are often referred to as augmented Krylov subspaces, since the basis of the space is augmented with the additional columns of P_1; see, e.g., [53, 17] and references given therein.

We also observe that in the semi-stable case, when $\Omega = 0$, we have $AP_1 = 0$, $A = A(P_1P_1^\top + P_2P_2^\top) = AP_2P_2^\top$ and

$$\mathcal{S}_k = \mathcal{R}(P_1) + \mathcal{K}_k(A, B) = \mathcal{K}_k(A, V_0), \tag{3.2}$$

the regular block Krylov subspace, augmented with $\mathcal{R}(P_1)$. Furthermore, when A is stable, P_1 is vacuous and the subspace in (3.1) reduces to $\mathcal{K}_k(A, B)$, so that the use of $\mathcal{S}_k$ as discussed above can be considered a true generalization of the standard case as originally proposed by Saad [49].

We remark here that the identities in (3.1), like all formulas developed in this paper, are only valid in exact arithmetic. In our numerical experiments, though, we have not observed any deterioration of our results due to round-off errors.

We mention as well that when linear independence in $\mathcal{S}_k$ is lost in the associated Arnoldi process, truncation using QR with partial column pivoting [18, p. 278] can be applied, as is done, e.g., in [10]. This limits any unnecessary growth in the rank of the approximate solutions. Alternatively, deflation can be used; see, e.g., [53, 17] and references given therein.

As is customary, we build an orthonormal basis of $\mathcal{S}_k$ with the Arnoldi process. At the kth step we have a matrix $V_k \in \mathbb{R}^{n\times\nu_k}$ whose columns span $\mathcal{S}_k$. With V_0 spanning the range of $[P_1, B]$ as previously described, the (block) Arnoldi process builds V_k and we obtain the (block) Arnoldi relation

$$AV_k = V_k\widehat{H}_k + \widehat{v}_{k+1}\widehat{r}_k^\top, \tag{3.3}$$

where $V_k^\top\widehat{v}_{k+1} = 0$, so that $\widehat{v}_{k+1}\widehat{r}_k^\top$ corresponds to the part of the range of AV_k orthogonal to $\mathcal{R}(V_k)$.

As described earlier, an approximate solution $X_k = V_kY_kV_k^\top$ is found by solving the projected Lyapunov equation (1.3). Both (1.1) and its projection (1.3) are singular and they are usually solved in the least squares sense.

To that end, let $V_k = [P_1, \widetilde{V}_k]$, $\widetilde{A}_{22} \equiv \widetilde{V}_k^\top A \widetilde{V}_k$, $\widetilde{A}_{12} \equiv P_1^\top A \widetilde{V}_k$, $\widetilde{B}_2 \equiv \widetilde{V}_k^\top B$, $Y_k = \begin{bmatrix} \widetilde{X}_{11} & \widetilde{X}_{12} \\ \widetilde{X}_{12}^\top & \widetilde{X}_{22} \end{bmatrix}$. Then, (1.3) has the equivalent formulation

$$\Omega \widetilde{X}_{11} + \widetilde{X}_{11}\Omega^\top + \widetilde{A}_{12}\widetilde{X}_{12}^\top + \widetilde{X}_{12}\widetilde{A}_{12}^\top + B_1 B_1^\top = 0, \tag{3.4a}$$

$$\Omega \widetilde{X}_{12} + \widetilde{X}_{12}\widetilde{A}_{22}^\top + \widetilde{A}_{12}\widetilde{X}_{22} + B_1 \widetilde{B}_2^\top = 0, \tag{3.4b}$$

$$\widetilde{A}_{22}\widetilde{X}_{22} + \widetilde{X}_{22}\widetilde{A}_{22}^\top + \widetilde{B}_2 \widetilde{B}_2^\top = 0. \tag{3.4c}$$

Note that (3.4a) is similar to (2.2a), both produced by the projection with P_1. The sLE can be both under- and over-determined, as shown in Section 2.2. It contains the only singular part of (1.1) and defines the freedom in $\widetilde{X}_{11}$. With the low-rank approximation $X_k = V_k Y_k V_k^\top$ after projection, $\widetilde{X}_{11}$ will likely be different from X_{11}, with terms in (3.4a) different from those in (2.2a). Equations (3.4b) and (3.4c) are similar to (2.2b) and (2.2c) respectively. They determine $\widetilde{X}_{12}$ and $\widetilde{X}_{22}$, respectively, the projected versions of X_{12} and X_{22}.

Algorithm 5.1 is a rendition of our proposed method as just described. One could consider other alternatives usually associated with nonsingular Lyapunov equations. We discuss these alternatives in the following remarks.

Remark 3.1. One may contemplate the use of However, as shown in the analysis in this section, this approach may not capture the part of the solution in $\mathcal{R}(P_1)$. We have performed a numerical experiment with RKS for Example 7.1, showing that in fact this approach is not competitive.

Remark 3.2 (Minimum Residual Solution). As we just described, with the space $\mathcal{S}_k$, we project the sLE (1.1) to a smaller sLE. For a large ν_k, we therefore may use the technique presented in Section 2, and we illustrate this with numerical experiments in Section 7. However, if ν_k is small enough, we can consider another alternative. Namely that the projected (singular) equation (1.3) be solved by any appropriate method. We may even apply "brute force", utilizing the Kronecker product $\otimes$ to produce

$$(I \otimes A_k + A_k \otimes I) v(Y_k) = -v(B_1 B_1^\top),$$

where $v(\cdot)$ stacks the columns of matrices. This can then be solved in the least squares or minimum residual sense by the QR or SVD factorizations. The least squares solution Y_k minimizes the residual of (1.3), without requiring any further assumptions on A or A_{22} as we do when we want to guarantee that the matrix $\widetilde{A}_{22}$ is stable, as discussed in Section 4 below. This minimum residual approach is summarized as Algorithm 5.2. We note however, that without separating or deflating the singular part in (1.3), as in (3.4a) in Algorithm 5.1 the original sLE and its projection are generally ill-posed. These problems are generally considered difficult to solve. In fact, one popular approach is the use of a regularization technique, such as Tikhonov regularization; see, e.g., [25]. Consequently, simply obtaining a minimum residual solution using Algorithm 5.2 is not expected to be efficient or accurate; see further a numerical experiment in Section 7.

Remark 3.3. The third alternative is to assess techniques similar to those presented by Lin and Simoncini [42]. However, the application of this approach directly assumes the nonsingularity of the sLE (1.1) as well as its projections. One could instead consider applying this technique to the nonsingular (3.4c). However, the associated projection of A_{22} lacks any particular structure such as sparsity and its stability has to be assumed. Since A_{22} is full (and this is why we usually keep it in sparse product form), using a RKS implies an expensive decompositions of $A_{22} - sI$ (for various shifts s).

4. Bounds on the Residuals and Stability Properties. In this section we present an *a posteriori* bound on the norm of the residual (1.2), and on other issues of convergence of the

methods described in the previous sections. In particular we address the question of whether $\widetilde{A}_{22}$ is stable if A_{22} is.

We present a result similar to [51, Proposition 5.1], which applies to our singular case.

PROPOSITION 4.1. *Consider a projection method defined by a matrix V_k to solve* (1.1)*, with approximation $X_k = V_k Y V_k^T$, and Arnoldi relation* (3.3)*. Let R_k be the residual* (1.2)*, and let $\widetilde{R}_k$ is the minimum residual in* (3.4a)*. Then,*

$$\|R_k\| \le \|\widetilde{R}_k\| + \|Y_k \widehat{r}_k\|, \quad \|R_k\|_F^2 = \|\widetilde{R}_k\|_F^2 + 2\|Y_k \widehat{r}_k\|_F^2. \tag{4.1}$$

Proof. Let $Q_1 \equiv V_k$ and extend it to an orthogonal matrix $Q = [Q_1, Q_2] \in \mathbb{R}^{n\times n}$. For the error or residual in the Krylov subspace projection method, from (3.3) and the fact that $\mathcal{R}(B) \subseteq \mathcal{R}(Q_1)$, we have

$$\begin{aligned} &Q_2^\top (AX_k + X_k A^\top + BB^\top) = Q_2^\top (AQ_1 Y_k Q_1^\top + Q_1 Y_k Q_1^\top A^\top) \\ &= Q_2^\top (Q_1 \widehat{H}_k^\top + \widehat{v}_{k+1}\widehat{r}_k^\top) Y_k Q_1^\top = Q_2^\top \widehat{v}_{k+1}\widehat{r}_k^\top Y_k Q_1^\top. \end{aligned}$$

We then have the residual

$$R_k = QQ^\top (AX_k + X_k A^\top + BB^\top) QQ^\top = Q \begin{bmatrix} \widetilde{R}_k & Y_k \widehat{r}_k \widehat{v}_{k+1}^\top Q_2 \\ Q_2^\top \widehat{v}_{k+1}\widehat{r}_k^\top Y_k & 0 \end{bmatrix} Q^\top,$$

from which we obtain (4.1). □

We can interpret the results in (4.1) as $\|Y_k \widehat{r}_k\|$ being the additional amount of residual created by the Krylov subspace projection method. In general, we do not know when $\|Y_k \widehat{r}_k\|$ is small *a priori*.

4.1. Stability of the projected matrix. As we mentioned earlier, when ν_k, the dimension of $\mathcal{S}_k$, is large, we may solve the projected sLE (1.3) using the reformulation (3.4).

With Ω possessing zero and the purely imaginary eigenvalues, the solvability of (3.4c) relies on the stability of $\widetilde{A}_{22} = \widetilde{V}_k^\top A \widetilde{V}_k$, i.e., on whether this matrix is stable.

In the rest of this section we discuss further whether the stability of the matrices $A_{22} = P_2^\top A P_2$ implies the stability of $\widetilde{A}_{22} = \widetilde{V}_k^\top A \widetilde{V}_k = (P_2 W_k)^\top A (P_2 W_k) = W_k^\top A_{22} W_k$. We investigate the effect of projection (by W_k) on the stability of A_{22}, or the distance of its spectrum from the imaginary axis. An appropriate tool is the *stability radius* [8, 26, 36, 37, 39, 54], which is nontrivial to estimate:

$$\psi(M) \equiv \min\{\|E\| : M + E \text{ is unstable}\} = \min_{\omega\in\mathbb{R}} \sigma_{\min}(M - \omega \imath I),$$

where $\sigma_{\min}$ is the smallest singular value.

Mirroring the Arnoldi relationship (3.3) for $A_{22}^\top$, let

$$A_{22}^\top W_k = W_k \widetilde{A}_{22}^\top + w_{k+1} s_k^\top, \tag{4.2}$$

with $w_{k+1} \in \mathcal{R}(\widetilde{W}_k)$. We have the following theorem on the inheritance of the stability of A_{22} by $\widetilde{A}_{22}$, modifying some techniques in [55, 56].

THEOREM 4.2. *Let $[W_k, \widetilde{W}_k]$ be orthogonal and $z_1 = \check{z}_1$ optimizes*

$$\psi(\widetilde{A}_{22}) = \min_{\omega\in\mathbb{R}} \min_{\|z_1\|=1} \left\| z_1^\top (\widetilde{A}_{22} - \imath\omega I) \right\|.$$

The stability of A_{22} is inherited by $\widetilde{A}_{22}$, with $\psi(\widetilde{A}_{22}) \geq \psi(A_{22}) - \|\check{z}_1^\top s_k\| > 0$, when

$$\psi(A_{22}) > \|\check{z}_1^\top s_k\|. \tag{4.3}$$

Proof. With $z = [z_1^\top, z_2^\top]^\top$, we have

$$\begin{aligned}
\psi(A_{22}) &= \psi([W_k, \widetilde{W}_k]^\top A_{22}[W_k, \widetilde{W}_k]) = \psi\left[\begin{array}{cc} W_k^\top A_{22} W_k & W_k^\top A_{22}\widetilde{W}_k \\ \widetilde{W}_k^\top A_{22} W_k & \widetilde{W}_k^\top A_{22}\widetilde{W}_k \end{array}\right] \\
&\leq \min_{\omega\in\mathbb{R}} \min_{\|z\|=1} \left\{ \left\| z^\top \left[\begin{array}{cc} \widetilde{A}_{22} - \imath\omega I & 0 \\ \widetilde{W}_k^\top A_{22} W_k & \widetilde{W}_k^\top A_{22}\widetilde{W}_k - \imath\omega I \end{array}\right] \right\| + \|z_1^\top W_k^\top A_{22}\widetilde{W}_k\| \right\} \\
&\leq \min_{\omega\in\mathbb{R}} \min_{\|z_1\|=1} \left\{ \left\| [z_1^\top, 0] \left[\begin{array}{cc} \widetilde{A}_{22} - \imath\omega I & 0 \\ \widetilde{W}_k^\top A_{22} W_k & \widetilde{W}_k^\top A_{22}\widetilde{W}_k - \imath\omega I \end{array}\right] \right\| + \|z_1^\top W_k^\top A_{22}\widetilde{W}_k\| \right\}, \\
&= \min_{\omega\in\mathbb{R}} \min_{\|z_1\|=1} \left\| z_1^\top (\widetilde{A}_{22} - \imath\omega I) \right\| + \|z_1^\top W_k^\top A_{22}\widetilde{W}_k\| \qquad (4.4) \\
&= \psi(\widetilde{A}_{22}) + \|\check{z}_1^\top W_k^\top A_{22}\widetilde{W}_k\|, \qquad (4.5)
\end{aligned}$$

where $z_1 = \check{z}_1$ optimizes the first term in (4.4) as assumed. From (4.2), $\|\check{z}_1^\top W_k^\top A_{22}\widetilde{W}_k\| = \|\check{z}_1^\top s_k\|$ and the inequalities (4.3) and (4.5) prove the result. □

A possible interpretation of the bound (4.3) comes from the following heuristic. When $\widetilde{A}_{22}$ is an increasingly good model of A_{22}, in terms of stability for the increasing k, z_1 will be dominated by the components at the top. When the Arnoldi residual s_k stagnates (otherwise it will be diminishing), it is usually dominated by components at the bottom. Consequently, $\|\check{z}_1^\top s_k\|$ diminishes with increasing k and (4.3) is eventually satisfied.

If condition (4.3) fails, the stability of A_{22} does not imply that of $\widetilde{A}_{22}$, although the latter may still hold for other reasons, e.g., A_{22} is passive i.e., the field of values $f(A_{22}) \equiv \{x^H A_{22} x : \|x\|^2 = 1, x \in \mathbb{C}^{n-\ell}\} \subseteq \mathbb{C}_-$, where X^H stands for the complex conjugate of x. Different scenarios occur, with A_{22} being passive but violating (4.3), and vice versa; see some examples in our numerical experiments in Section 7, where we use the techniques in [35] to compute the field of values.

Remark 4.1 (Solvability of Projected Equations). Apart from the additional assumptions on A_{22} in Theorem 4.1, there may well be other circumstances under which (3.4c) is solvable. Without any assumptions on A or A_{22}, the minimum residual approach in Remark 3.1 may also be applied. In the literature on projection methods, it is generally assumed that either (i) A_{22} is passive or (ii) the associated projected equation is solvable; see, e.g., the survey [52], and references therein. Note that (ii) is much stronger than the stability of A_{22} and is not easy to test or characterize. We have made a slight progress in avoiding these assumptions but further research is warranted on the treatment and solvability of projected equations.

Remark 4.2 (Special Case). When $\Omega = 0$, then we may again choose the free variable $X_{11} = 0$ for the minimum norm least squares solution, yielding the residual $R \equiv A_{12}X_{12}^\top + X_{12}A_{12}^\top + B_1 B_1^\top$. Equation (2.2c) yields the unique symmetric positive definite X_{22} and $X_{12} = -(A_{12}X_{22} + B_1 B_2^\top)A_{22}^{-\top}$. Consequently for (1.1), the corresponding least squares solution is $X = P_1 X_{12} P_2^\top + P_2 X_{12}^\top P_1^\top + P_2 X_{22} P_2^\top$.

5. Algorithms for large sLEs. We summarize in this section the two approaches we consider with the space $\mathcal{S}_k$ presented in Section 3, namely the reformulation (3.4) and the projected sLE in (1.3) respectively. In each case, we indicate the flop count at each step, for which we assume that there are at most c nonzero elements on each row or column of A.

The first algorithm we describe here is based on the reformulation (3.4), in conjunction with the use of the space $\mathcal{S}_k$, and which is guaranteed to work as long as the projected matrix $\widetilde{A}_{22}$ is stable, though it also works in other cases, as we mentioned in Remark 4.1. See also Section 7 where this is used for several examples.

ALGORITHM 5.1 (with stable $\widetilde{A}_{22}$).

1. For a given k, P_1 and $V_k = [P_1, \widetilde{V}_k]$ respectively with orthonormal columns, compute $\Omega = P_1^\top A P_1$, $\widetilde{A}_{12} = P_1^\top A \widetilde{V}_k$, $\widetilde{A}_{22} = \widetilde{V}_k^\top A \widetilde{V}_k$, $B_1 = P_1^\top B$ and $\widetilde{B}_2 = \widetilde{V}_k^\top B$; $2[(\nu_k - \ell)^2 + c(\nu_k - \ell) + \ell\nu_k + lp]n$ flops.
2. Solve the Lyapunov equation (3.4c) for $\widetilde{X}_{22}$; $2(\nu_k - \ell)^3$ flops.
3. Solve the Sylvester equation (3.4b) for $\widetilde{X}_{12}$; $l\nu_k(\nu_k - \ell)$ flops.
4. Solve (3.4a) in the least squares sense for $\widetilde{X}_{11}$; $O(\ell^6)$ flops.
5. Estimate ρ_k as in (4.1).
6. If ρ_k is acceptably small, exit; otherwise $k \leftarrow k+1$, increase the dimension of V_k by the Arnoldi process, with truncation by the QR process with partial column pivoting, and go back to Step 1; $(14\nu_k^2 n + 20\nu_k^3/3)$ flops.
7. Let the approximate solution $X_k = V_k Y_k V_k^\top = P_1 \widetilde{X}_{11} P_1^\top + P_1 \widetilde{X}_{12} \widetilde{V}_k^\top + \widetilde{V}_k \widetilde{X}_{12}^\top P_1^\top + \widetilde{V}_k \widetilde{X}_{22} \widetilde{V}_k^\top$. From the SVD of $Y_k = U_k \Sigma_k U_k^\top + \widetilde{U}_k \widetilde{\Sigma}_k \widetilde{U}_k^\top$, where $\|\widetilde{\Sigma}_k\|$ is smaller than some given tolerance, truncate Y_k of negligible components to $U_k \Sigma_k U_k^\top$ and form the truncated approximate solution $\widetilde{X}_k = \widetilde{V}_k \widetilde{Y}_k \widetilde{V}_k^\top$ with $\widetilde{V}_k \equiv V_k \widetilde{U}_k$ and $\widetilde{Y}_k \equiv \widetilde{\Sigma}_k$; $O(\nu_k^3)$ flops.

The total flop count for an iteration in the Krylov subspace in Algorithm 5.1 is $2[8\nu_k^2 + (c + \ell)\nu_k + \ell p]n$ flops, ignoring $O(\ell)$ or other constant terms, while assuming $\nu_k \ll n$. When computing V_{k+1} in Step 6, for a given tolerance, the QR factorization with partial column pivoting is applied to $[V_k, AV_k]$. This controls any unnecessary growth of the width ν_k of V_k. A final truncation is applied in Step 7, which limits further the dimension of the approximate solution $\widetilde{X}_k$. Step 7 is applied *a posteriori* after the growth of the residual norm ρ_k is judged to be acceptably small in Step 6. This is standard for Lyapunov solvers with Krylov projection; see, e.g., [38]. The determination of the stability of $\widetilde{A}_{22}$ is expensive, whether based on the inequality (4.3) or the passivity of A. We have not included these costs in the flop counts.

ALGORITHM 5.2.

1. For a given k, V_k with orthonormal columns, compute $A_k = V_k^\top A V_k$ and $B_k = V_k^\top B$; $2\nu_k(c + p)n$ flops.
2. Solve the projected singular Lyapunov equation (1.3) for Y_k in the minimum residual sense; $\nu_k^6/3$ flops.
3. Estimate the norm of the residual contributed by projection $\rho_k \equiv \|Y_k \widehat{r}_k\|$ as in (4.1); $O(\nu_k^3)$ flops.
4. If ρ_k is acceptably small, go to Step 5; otherwise $k \leftarrow k+1$, increase the dimension of V_k by the Arnoldi process, with truncation by the QR process with partial column pivoting, and go back to Step 1; $(14\nu_k^2 n + 20\nu_k^3/3)$ flops.
5. From the SVD of $Y_k = U_k \Sigma_k U_k^\top + \widetilde{U}_k \widetilde{\Sigma}_k \widetilde{U}_k^\top$, where $\|\widetilde{\Sigma}_k\|$ is smaller than some given tolerance, truncate Y_k of negligible components to $U_k \Sigma_k U_k^\top$ and form the truncated approximate solution $\widetilde{X}_k = \widetilde{V}_k \widetilde{Y}_k \widetilde{V}_k^\top$ with $\widetilde{V}_k \equiv V_k \widetilde{U}_k$ and $\widetilde{Y}_k \equiv \widetilde{\Sigma}_k$; $O(\nu_k^3)$ flops.

The total flop count for an iteration in the Krylov subspace in Algorithm 5.2 is $2\nu_k(7\nu_k + cp)n$ flops, ignoring $O(\ell)$ or other constant terms, while assuming $\nu_k \ll n$.

When computing V_{k+1} in Step 4, for a given tolerance, the QR factorization with partial column pivoting is applied to $[V_k, AV_k]$ to form V_{k+1}, eliminating the components in $[V_k, \widehat{v}_{k+1}]$

from (3.3) which are nearly linear dependent, as in Step 6 in Algorithm 5.1. A final truncation is applied in Step 5, as in Step 7 in Algorithm 5.1. Step 5 is applied *a posteriori* after the growth of the residual norm ρ_k is judged to be acceptably small in Step 4.

We are not aware of any way to know *a priori* when to stop the iterations earlier, given the possible lower rank of the solution. We mention however, that in other contexts, there has been experimentation (but no theory), with truncating different elements such as the direction matrix in (block) conjugate gradients, throughout the iterations, so that the subspaces do not grow as much; see, e.g., [2, 40, 50]. To keep things focused, we chose not to explore this alternative in the current paper, but it could be considered in a larger study.

We end this section mentioning that for either algorithm to work, a reasonable accurate P_1 is required. In many practical cases, as illustrated by the numerical examples in Section 7 this is perfectly feasible. We also mention that, like all iterative methods, if the algorithm takes too many iterations to converge to certain tolerance, one may run out of memory. A lower accuracy may be achieved by the approximate solutions for the same memory usage.

6. Singular Stein Equations. In this section, we show that we can use an approach very similar to those presented in sections 2 and 3 for the solution of the singular Lyapunov equation (1.1) for the singular semi-stable Stein equation

$$-X + AXA^\top + BB^\top = 0, \tag{6.1}$$

with $\Lambda(A)$ lying inside the unit circle except for a few semi-simple unimodular eigenvalues. With A written as in (2.1), we have the following result.

THEOREM 6.1 (Singular Stein Equations). *Given the matrix A and the orthogonal matrix $P = [P_1, P_2]$ such that one can write* (2.1)*, the sSE* (6.1) *is equivalent to*

$$-X_{11} + \Omega X_{11}\Omega^\top + A_{12}X_{12}^\top\Omega^\top + \Omega X_{12}A_{12}^\top + A_{12}X_{22}A_{12}^\top + B_1B_1^\top = 0, \tag{6.2a}$$
$$-X_{12} + \Omega X_{12}A_{22}^\top + A_{12}X_{22}A_{22}^\top + B_1B_2^\top = 0, \tag{6.2b}$$
$$-X_{22} + A_{22}X_{22}A_{22}^\top + B_2B_2^\top = 0, \tag{6.2c}$$

where $X_{ij} \equiv P_i^\top X P_j$ and $B_i \equiv P_i^\top B$ $(i, j = 1, 2)$, analogous to (2.2)*. Consequently, X_{22} is uniquely solved from the nonsingular Stein equation* (6.2c) *and X_{12} can be solved from the nonsingular generalized Sylvester equation* (6.2b)*. Then* (6.2a) *is solved in the least squares sense, obtaining X_{11}. Then $X = \sum_{i,j} P_i X_{ij} P_j^\top$ is the least squares solution to the sSE* (6.1)*, with its residual identical in norm to that of* (6.2a) *for the corresponding least squares solutions.*

Proof. Writing A as in (2.1), the Stain equation (6.1) is equivalent to

$$-\begin{bmatrix} X_{11} & X_{12} \\ X_{12}^\top & X_{22} \end{bmatrix} + \begin{bmatrix} \Omega & A_{12} \\ 0 & A_{22} \end{bmatrix}\begin{bmatrix} X_{11} & X_{12} \\ X_{12}^\top & X_{22} \end{bmatrix}\begin{bmatrix} \Omega^\top & 0 \\ A_{12}^\top & A_{22}^\top \end{bmatrix} + \begin{bmatrix} B_1B_1^\top & B_1B_2^\top \\ B_2B_1^\top & B_2B_2^\top \end{bmatrix} = 0,$$

and in turn to (6.2), where A_{22} is stable (with eigenvalues inside the unit circle for Stein equations) and Ω contains a few semi-simple unimodular eigenvalues. Therefore, (6.2b) and (6.2c) are nonsingular and the inconsistency in (6.1) appears only in (6.2a), whose residuals have the same norm for the corresponding least squares solutions. □

As in Section 3 for sLEs, the appropriate subspaces can be chosen to be (3.1) in this case as well. The algorithms for sSEs, are very similar to those for the sLEs, and are therefore omitted.

On inconsistency, for an illustration with $\Omega = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ where $\alpha^2 + \beta^2 = 1$, (6.2a) has the

form

$$-\begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} + \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} + \begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix} = 0$$
$$\Leftrightarrow \begin{bmatrix} 2\alpha\beta & \beta^2 \\ 2\alpha\beta & \alpha^2 \\ 2\alpha\beta & \alpha^2 \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{22} - x_{11} \end{bmatrix} + \begin{bmatrix} q_{11} \\ -q_{22} \\ -q_{12} \end{bmatrix} = 0.$$

The generically inconsistency of the equation and the extra freedom in $x_{22} - x_{11}$ are evident.

When $\Omega = -I$ in the semi-stable case, we choose the free parameter $X_{11} = 0$ for the least squares solution for the generically inconsistent (6.2a), then the minimum residual equals $-A_{12}X_{12}^\top - X_{12}A_{12}^\top + A_{12}X_{22}A_{12}^\top + B_1B_1^\top$, $X_{12} = (A_{12}X_{22}A_{22}^\top + B_1B_2^\top)(I + A_{22}^\top)^{-1}$ and X_{22} can be solved from the standard Stein equation (6.2c).

When A is symmetric, Ω and A_{22} are symmetric and $A_{12} = 0$, with (6.2) becoming

$$-X_{11} + \Omega X_{11}\Omega + B_1B_1^\top = 0, \quad -X_{12} + \Omega X_{12}A_{22} + B_1B_2^\top = 0, \quad -X_{22} + A_{22}X_{22}A_{22} + B_2B_2^\top = 0,$$

which are decoupled. When $\Omega = -I$, the equations reduce even further to

$$-X_{12}(I + A_{22}) + B_1B_2^\top = 0, \quad -X_{22} + A_{22}X_{22}A_{22} + B_2B_2^\top = 0,$$

with $X_{11} = 0$ for the least squares solution and $B_1B_1^\top$ being the minimum residual.

7. Numerical Examples. We present three numerical experiments illustrating the approaches presented in Sections 2 and 3 for the solution of the singular Lyapunov equation (1.1), namely for medium and large n, respectively. The first two examples correspond to large n, of $n = 10000$ and $n = 10201$, and in both cases, we have $\Omega = 0$, i.e., the semi-stable case, while in the third example, we have $n = 1000$ and $\Omega \neq 0$ with two nonzero eigenvalues on the imaginary axis. In addition to reporting on convergence results, we look at the eigenvalues, or the field of values of the appropriate matrices, and discuss when the projected matrix $\tilde{A}_{22}$ maintains the stability property of A_{22}. For Example 7.1, we have also included some numerics by the approaches discussed in Remarks 3.1–3.2, indicating their lack of competitiveness for this singular case.

As we described in Section 2, for medium-scale sLEs, we proceed as in Algorithm 2.1 we solve first (2.2c), and then in turn (2.2b). Both equations are solved using the standard Bartels-Stewart algorithm [1], implemented by the MATLAB command `lyap` [43]. The singular (2.2a) is solved in the least squares sense by applying Kronecker products and the SVD [18].

Solving large-scale sLEs by the projection method, except when stated otherwise, we proceed as in Algorithm 5.1: we solve the projected Lyapunov equation (1.3) as in (3.4), all in smaller dimensions. Equation (3.4c) and subsequently (3.4b) are solved using the MATLAB command `lyap` [43]. The singular (3.4a) is solved as described above for the medium-size problem.

All numerical experiments have been carried out on a computer with an Intel Core i7-4721HQ CPU @2.3GHZ, 16GB memory and Windows 10 operating system. For the Krylov subspace projection method for large-scale sLEs, the tolerance for the truncation in the QR decomposition with column pivoting is 10^{-10}. The numerical rank with the tolerance $\varepsilon = 10^{-16}$ is denoted by rank_ε, in Step 5 of Algorithm 5.1.

For the computation of the low-dimensional $\mathcal{R}(P_1)$, inverse iteration was applied to A. In our examples this was obtained without any complications. As we shall see, even when A_{22} or $\widetilde{A}_{22}$ have eigenvalues near the imaginary axis, our method works accurately. The solution of the stable Lyapunov equations (2.2c) or (3.4c) dominate in terms of execution time.

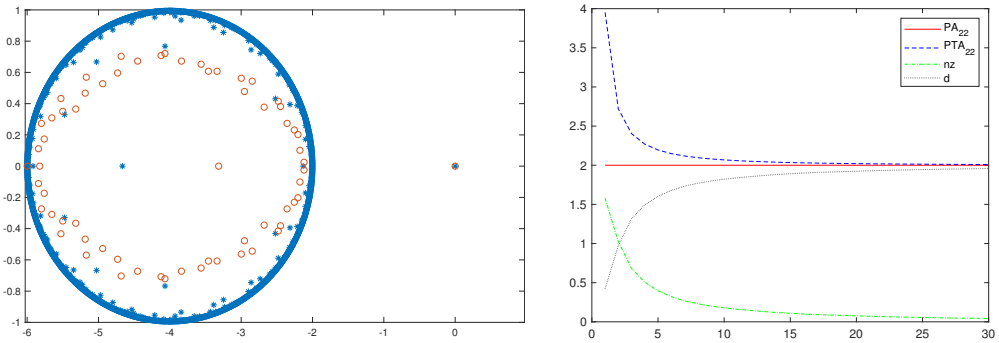
To verify (4.5) for the large-scale Examples 7.1 and 7.2, we estimate the stability radii $\psi(A_{22})$ and $\psi(\widetilde{A}_{22})$ using the eigenvalue optimization software `Eigopt` [44] (with a tolerance for accuracy

of 10^{-12}), which implements the methods in [7, 45]. The field of values of A_{22} has been computed using the command `fov` in `chebfun` [11], which implements the techniques in [35].

EXAMPLE 7.1. [33, Example 1] We consider the matrix

$$A = -\begin{bmatrix} 4 & 1-\delta & 0 & \cdots & 0 & 1 \\ 1+\delta & 4 & 1-\delta & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \ddots & \ddots & 1-\delta \\ 1 & 0 & \cdots & 0 & 1+\delta & 4 \end{bmatrix}.$$

with $\delta = 0.5$ and $n = 10000$, and we force $\dim \mathcal{N}(A) = \ell$ by setting the last ℓ rows of A to be zeros. The entries in B were uniformly distributed on $[0, 1]$. We compare the spectra $\Lambda(A)$ and $\Lambda(V_k^\top A V_k)$ with $\ell = 1$ and $k = 21$ in Figure 7.1(a). The figure shows the stability property of A_{22} is passed on to $\widetilde{A}_{22}$. Notice the singularity of A_{11} at the origin. The spectral plots for $\ell = 3, 6$ show similar behavior. We did not compute the plot for the field of values of A_{22} because of the large value of n, but the largest eigenvalue of the symmetric part $(A_{22} + A_{22}^\top)/2$ equals 0.0818, implying that A_{22} is not passive. We also display $\mathrm{PA}_{22} = \psi(A_{22})$, $\mathrm{PTA}_{22} = \psi(\widetilde{A}_{22})$, $\mathrm{nz} = \|\check{z}_1^\top s_k\|$ and $\mathrm{d} = \psi(A_{22}) - \|\check{z}_1^\top s_k\|$ for $k = 1, \ldots, 30$ in Figure 7.1(b). The computation of $\psi(A_{22})$ and $\psi(\widetilde{A}_{22})$ (for each k and ℓ) was performed with `Eigopt` as mentioned above. To verify (4.5) and (4.3) for stable A_{22} and $\widetilde{A}_{22}$, all the displayed quantities have to be positive, and this holds starting with $k = 1$. Thus, the stability of A_{22} passes to $\widetilde{A}_{22}$ after projection and the projected Lyapunov equation (3.4c) is uniquely solvable.



(a) $\Lambda(A)$ (blue stars) and $\Lambda(V_k^\top A V_k)$ (red circles) (b) $\psi(A_{22})$, $\psi(\widetilde{A}_{22})$, $\|\check{z}_1^\top s_k\|$ and $\psi(A_{22}) - \|\check{z}_1^\top s_k\|$

Fig. 7.1: Example 7.1. Spectra and stability radii

For different values of $\dim \mathcal{N}(A)$, $\operatorname{rank} B$, and $k = 30$, we display in Table 7.1 the corresponding execution time (in seconds) required for the projection method, ν_k (the number of columns of V_k), $\operatorname{rank}_\varepsilon(X_k)$ and the normalized residual $\mu_k \equiv \rho_k/(2\|A\|\|Y_k\| + \|B\|^2)$ with $\rho_k \equiv \|Y_k \widehat{r}_k\|$ from (4.1). The computation of $P_1 \in \mathbb{R}^{n\times\ell}$ (with $\ell = 1, 3, 6$) which spans $\mathcal{N}(A)$ took between 0.0469 sec. and 0.0781 sec.

We also present the graphs of the normalized residual μ_k in Figure 7.2 for $\dim \mathcal{N}(A) = 1$ and $\operatorname{rank}(B) = 3, 7$ and 11. The graphs for different values of $\operatorname{rank}(B)$ are visually identical, and

$\dim(\mathcal{N}(A))$	$\text{rank}(B)$	cputime	ν_k	$\text{rank}_\varepsilon(X_k)$	N_k
1	3	4.78	91	26	1.40×10^{-16}
	7	8.17	197	57	9.42×10^{-17}
	11	14.73	331	89	8.98×10^{-17}
3	3	6.31	93	28	2.40×10^{-16}
	7	9.19	213	59	1.40×10^{-16}
	11	16.33	333	91	1.26×10^{-16}
6	3	7.51	96	30	3.10×10^{-16}
	7	11.15	216	62	2.20×10^{-16}
	11	16.70	336	94	1.80×10^{-16}

Table 7.1: Example 7.1. Projection Methods at $k = 30$

so are those for $\dim \mathcal{N}(A) = 3, 6$. A tolerance of $O(10^{-16})$ in terms of normalized residual μ_k is achieved after 30 iterations. The table and figure illustrate the feasibility and efficiency of the projection method.

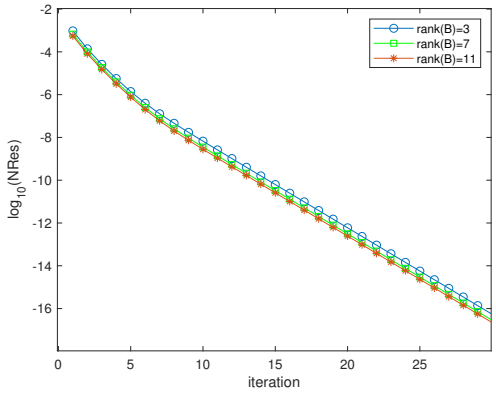


Fig. 7.2: Example 7.1. Normalized residuals for $\dim(\mathcal{N}(A)) = 1$, $\text{rank}(B) = 3, 7$ and 11.

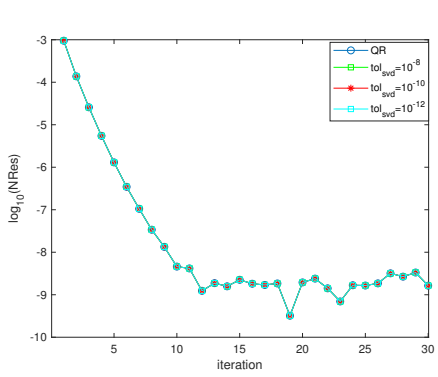
We consider now other alternatives to our proposed method, as described in Remarks 3.1–3.3. We present results of three additional experiments. In these experiments, $\dim(\mathcal{N}(A)) = 1$ and B is chosen randomly with its rank being 3. The tolerance "tol_{svd}" controls the truncation of the corresponding SVD in the least squares solution. The shifts for the rational Krylov subspaces are chosen as in [47].

In the first experiment, the RKSs generated by A, as discussed in Remark 3.1, is applied to the sLE (1.1) and the projected equations are solved in the least squares sense (by the SVD [18]). The normalized residuals are at least of $O(10^3)$, showing the subspaces are inappropriate.

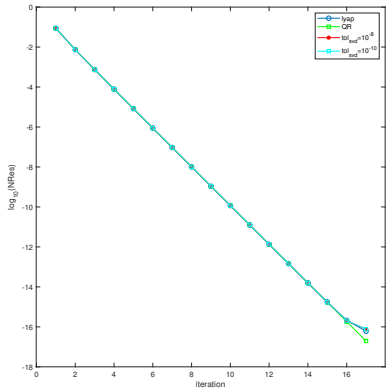
On the second experiment summarized on the left of Figure 7.3, the RKS approach is applied to (1.3), as in Algorithm 5.2 or Remark 3.2. The least squares solution is realized by the backslash \ in MATLAB [43] (labelled by "QR") and the SVD [18] (with three difference values of the tolerance "tol_{svd}" for truncating small singular values). The CPU time required is approximately 2 sec. for 8 iterations, achieving an accuracy of $O(10^{-9})$, similar to that for Algorithm 5.1 at the similar stage of iteration. However, the accuracy cannot be improved much with further iterations.

On the third experiment on the right of Figure 7.3, the RKS approach is applied only to

(2.2c), as suggested in Remark 3.3. The projected equation is then solved by the command `lyap` in MATLAB [43] (assuming solvability, labelled by "lyap" in the figure), as well as in the minimum residual sense by the backslash \ in MATLAB [43] (labelled by "QR") and the SVD [18] (with two difference values of the tolerance "$\mathrm{tol}_{\mathrm{svd}}$" for truncating small singular values). Similar accuracy as from Algorithm 5.1 is achieved while requiring more than 200sec. for 30 iterations, exceeding 40 times the CPU time required by Algorithm 5.1.



(a) Normalized residuals for Algorithm 5.2



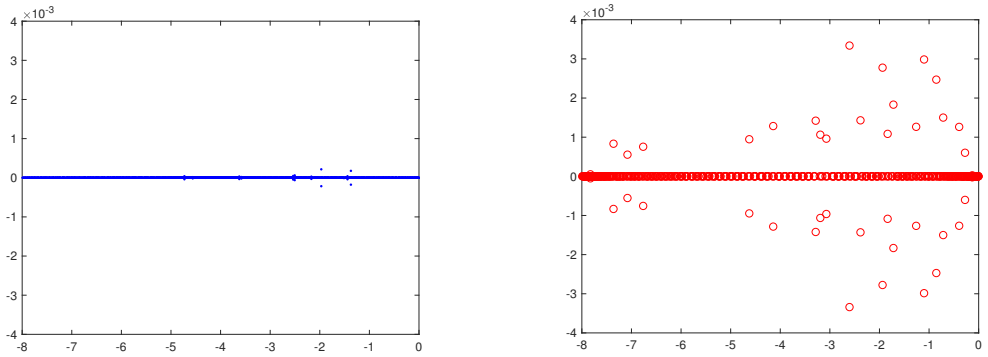
(b) Normalized residuals for RKS on (2.2c)

Fig. 7.3: Example 7.1. Rational Krylov and minimum residual methods

EXAMPLE 7.2. We consider the PDE operator $-\Delta u + \vec{a}\nabla u$ with Neumann boundary conditions. Using the finite element method with triangular elements and quadratic basis functions, we discretize to obtain the matrix A [15, Chapters 3 and 4]. It is well-known that A has a semi-stable spectrum spreading out from the origin and $\dim(\mathcal{N}(A)) = 1$. We experiment with (i) $n = 2601$ and $\vec{a} = 1$ in 1D and (ii) $n = 10201$ and $\vec{a} = (1, 2)$ in 2D. We chose B randomly with rank $5, 9$ and 11. The Δu term dominates thus A is not far from being symmetric and $\Lambda(A_{22}) \subset \mathbb{C}_-$ (A_{22} is stable) but situated close to the origin. We expect A_{22} in (2.2c) and $\widetilde{A}_{22}$ in (3.4c) to be ill-conditioned, with many eigenvalues small in magnitude. This is a good test for our method.

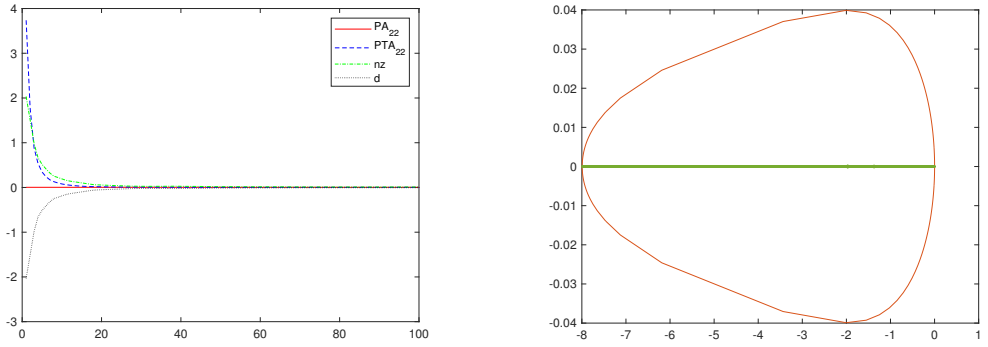
In Figure 7.4, we compare the spectra $\Lambda(A)$ and $\Lambda(V_k^\top A V_k)$ (as we did in Figure 7.1(a) for Example 7.1) for $n = 2601$ with $\ell = 1$ and $k = 100$ (respectively indicated with blue stars and red circles). The spectra are not potted together to avoid fuzziness. They are away from the negative real axis with negligible imaginary parts. Obviously, the gap between $\Lambda(A_{22})$ and the imaginary axis is so small, that (4.3) is not satisfied, as indicated by $\mathrm{d} < 0$ in Figure 7.5(a) (with the same symbols as in Figure 7.1(b)). However, the field of values of A_{22} in Figure 7.5(b) with the right vertical tangent just on the left of the origin, given by the maximum eigenvalue of $(A_{22} + A_{22}^\top)/2$ of -0.31×10^{-2} and -0.81×10^{-3} respectively for $n = 2601, 10201$. This illustrates that A_{22} is passive. (Only the field of values of A_{22} for $n = 2601$ is displayed, that of $n = 10201$ would be too expensive to compute.) This implies the stability of $\widetilde{A}_{22}$ and the unique solvability of (3.4c). The graphs for $n = 10201$ in 2D shows similar behaviour.

We iterate to $k = 100, 150$ steps respectively for $n = 2601, 10201$ as summarized in Table 7.2. Because of its spreading spectrum, $\widetilde{A}_{22}$ is ill-conditioned and the projection method converges slowly. Figure 7.6 shows the relative residual norm μ_k respectively for $n = 2601, 10201$. More accurate approximate solutions may be obtained with larger k. Iterations with a larger rank (B)



(a) $\Lambda(A_{22})$ (b) $\Lambda(\widetilde{A}_{22})$

Fig. 7.4: Example 7.2. Spectra



(a) $\psi(A_{22})$, $\psi(\widetilde{A}_{22})$, $\|\tilde{z}_1^\top s_k\|$ and $\psi(A_{22}) - \|\tilde{z}_1^\top s_k\|$ (b) Field of Values and Spectrum of A_{22}

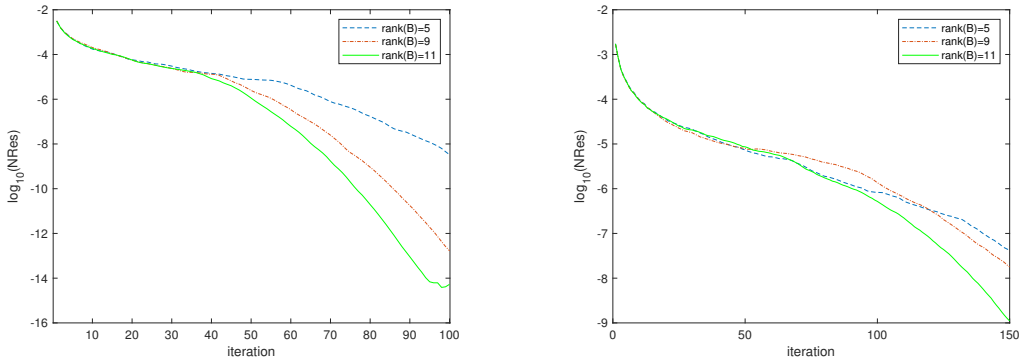
Fig. 7.5: Example 7.2. Stability radii, field of values and spectrum

converge in fewer iterations with bigger Krylov subspaces but require more execution time.

n	rank(B)	cputime	ν_k	rank(X_k)	μ_k
2601	5	77.45	501	89	3.1×10^{-9}
	9	182.20	901	146	1.9×10^{-13}
	11	265.08	1097	172	5.9×10^{-16}
10201	5	186.42	751	99	1.3×10^{-6}
	9	603.03	1351	174	1.3×10^{-8}
	11	867.70	1651	194	1.4×10^{-9}

Table 7.2: Example 7.2. Projection Methods at $(n, k) = (2601, 100)$ and $(10201, 150)$

EXAMPLE 7.3. We construct A from the matrix of Example 7.1 with $n = 1000$ and $\delta = 0.5$,



(a) $n = 2601$ (b) $n = 10201$

Fig. 7.6: Example 7.2. Normalized residuals

except for zeros in the last two rows and columns as well as the diagonal block $\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$ at the lower right corner. Consequently, A has a pair of pure imaginary eigenvalues $\pm 2i$, in addition to the other stable eigenvalues, as shown in Figure 7.7. The elements in the matrix $B \in \mathbb{R}^{1000\times 5}$ are uniformly distributed in (0,1). The MATLAB command `lyap` [43] was applied to (2.2c) and (2.2b) to obtain X_{22} and X_{12}, respectively, with negligible errors. For the solution of (2.2a), we applied Kronecker products to produce the singular linear equation $M\mathrm{vec}(X_{11}) = f \equiv -\mathrm{vec}(A_{12}X_{12}^\top + X_{12}A_{12}^\top + B_1B_1^\top)$, where $\mathrm{vec}(\cdot)$ stacks the columns of a given matrix. We then obtained the SVD of the coefficient matrix, i.e., $M = U_1DV_1^\top$ and constructed the least squares solution $\mathrm{vec}(X_{11}) = V_1D^{-1}U_1^\top f$ [18]. We ignore any singular values less than $\epsilon = 10^{-14}\sigma_1$ in D, in order to avoid misreading the singularity of the Lyapunov operator. For the residual $r \equiv f - M\mathrm{vec}(X_{11})$, we have $\|r\| = 1.4232$ and the norm of the residual of the corresponding normal equation $\|M^\top r\| = 1.26\times 10^{-15}$. This calculation took a total of 2.99 seconds of execution time. This indicates that the equations in (2.2) are solved efficiently up to machine accuracy.

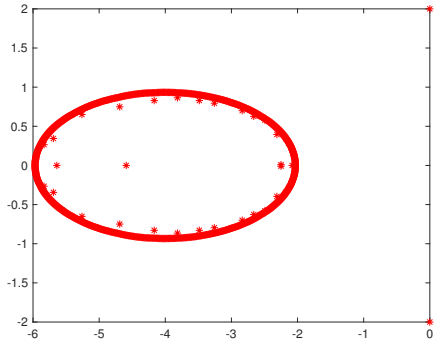


Fig. 7.7: Example 7.3. $\Lambda(A)$

8. Conclusions. We have analyzed the singular Lyapunov and Stein equations. We have shown that the singular equations can be decoupled into a singular part, and two uniquely solvable equations, from which the least squares solutions can be derived for small problems. For large-scale equations, the appropriate Krylov subspaces have been identified for the corresponding projection method. Illustrative numerical examples of sLEs have been presented indicating the effectiveness of the proposed method.

Acknowledgements. Part of the work was carried out when the first and third authors respectively visited the School of Mathematical Sciences at Fudan University in 2018–19 and the Department of Mathematics at Temple University in 2016. The authors thank the anonymous referees for their questions and suggestions which led to improvements in our presentation.

REFERENCES

[1] R.H. Bartels, and G.W. stewart. Solution of the matrix equation $AX+XB=C$, *Comm. ACM*, 15 (1972) 820–826.
[2] P. Benner, and T. Breiten. Low rank methods for a class of generalized Lyapunov equations and related issues. *Numer. Math.*, 124 (2013) 441–470.
[3] P. Benner, J.-R. Li, and T. Penzl. Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratric optimal control problems, *Numer. Linear Algebra Applic.*, 15 (2008) 755–777.
[4] D.S. Bernstein, and S.P. Bhat. Lyapunov stability, semistability, and asymptotic stability of matrix second-order systems, *J. Mech. Design*, 117B (1995) 145–153.
[5] S.P. Bhat, and D.S. Bernstein. Nontangency-based Lyapunov tests for convergence and stability in systems having a continuum of equilibria, *SIAM J. Control Optim.*, 42 (2003) 1745–1775.
[6] S.P. Bhat, and D.S. Bernstein, Arc-length-based Lyapunov tests for convergence and stability with applications to systems having a continuum of equilibria, *Math. Control, Signals & Systems*, 22 (2010) 155–184.
[7] L. Breiman, and A. Cutler. A deterministic algorithm for global optimization, *Math. Program.*, 58 (1993) 179–199.
[8] R. Byers. A bisection method for measuring the distance of a stable matrix to the unstable matrices, *SIAM J. Sci. Stat. Comput.*, 9 (1988) 875–881.
[9] S.L. Campbell, and N.J. Rose. Singular perturbation of autonomous linear systems, *SIAM J. Math. Anal.*, 10 (1979) 542–551.
[10] E.K.-W. Chu, T. Li, W.-W. Lin, and C.-Y. Weng. Solving large-scale continuous-time algebraic Riccati equations by doubling, *J. Comput. Appl. Math.*, 237 (2013) 373–383.
[11] T. A. Driscoll, N. Hale, and L. N. Trefethen, eds.. *Chebfun Guide*, Pafnuty Publications, Oxford, 2014.
[12] 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) 1875–1898.
[13] V. Druskin, and V. Simoncini. Adaptive rational Krylov subspaces for large-scale dynamical systems, *Syst. Control Letters*, 60 (2011) 546–560.
[14] A. El Guennouni, K. Jbilou, and A.J. Riquet. Block Krylov subspace methods for solving large Sylvester equations *Numer. Algor.* 29 (2002) 75–96.
[15] H.C. Elman, D.J. Silvester, and A.J. Wathen. *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics*, Oxford University Press, Oxford, 2005.
[16] A. Frommer, K. Lund, and D.B. Szyld. Block Krylov subspace methods for functions of matrices, *Electr. Trans. Numer. Anal.*, 47 (2017) 100–126.
[17] A. Gaul, M.H. Gutknecht, J. Liesen, and R. Nabben. A framework for deflated and augmented Krylov subspace methods, *SIAM J. Matrix Anal. Appl.*, 34 (2013) 495–518.
[18] G.H. Golub, and C.F. Van Loan. *Matrix Computations*, 4th ed., Johns Hopkins University Press, Baltimore, 2013.
[19] C. Gotsman, and S. Toledo. On the computation of null spaces of sparse rectangular matrices, *SIAM J. Matrix Anal. Appl.*, 30 (2008) 445–463.
[20] M.H. Gutknecht. *Block Krylov space methods for linear systems with multiple right-hand sides: An introduction*, in Mod. Math. Model. Methods Algorithms Real World Syst., A.H. Siddiqi, I.S. Duff, and O. Christensen, eds., New Delhi, 2007, Anamaya, pp. 420–447.
[21] W.M. Haddad, V.S. Chellaboina, S.P. Bhat, and D.S. Bernstein. Modeling and analysis of mass-action kinetics, *IEEE Control Syst. Magazine*, 29 (2009) 60–78.
[22] W.M. Haddad, V.S. Chellaboina, and Q. Hui. *Nonnegative and Compartmental Dynamical Systems*, Princeton University Press, Princeton, 2010.

[23] W.M. Haddad, V.S. Chellaboina, and S.G. Nersesov. *Thermodynamics: A Dynamical Systems Approach*, Princeton University Press, Princeton, 2005.
[24] W.M. Haddad, Q. Hui, and V.S. Chellaboina. H_2 optimal semistable control for linear dynamical systems: an LMI approach, *J. Franklin Inst.*, 348 (2011) 2898–2910.
[25] P.C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion, 1st ed.*, SIAM, Philadelphia, 1998.
[26] G. Hu and F.J. Davison. Real controllability/stabilizability radius of LTI systems, *IEEE Trans. Automat. Control*, 49 (2004) 254–257.
[27] D.Y. Hu, and L. Reichel. Krylov-subspace methods for the Sylvester equation, *Linear Algebra Appl.*, 172 (1992) 283–313.
[28] Q. Hui. Optimal semistable control for continuous-time linear systems, *Syst. Control Letters*, 60 (2011) 278–284.
[29] Q. Hui, and J.M. Berg. Semistability theory for spatially distributed systems, *Syst. Control Letters*, 62 (2013) 862–870.
[30] Q. Hui, and W.M. Haddad. Distributed nonlinear control algorithms for network consensus, *Automatica*, 44 (2008) 2375–2381.
[31] Q. Hui, and W.M. Haddad. Finite-time semistability and consensus for nonlinear dynamical networks, *IEEE Trans. Autom. Control*, 53 (2008) 1887–1900.
[32] D.J. Inman. *Vibration with Control*, John Wiley & Sons, Chichester, West Sussex, England, 2006.
[33] K. Jbilou. Block Krylov subspace methods for large algebraic Riccati equations, *Numer. Algor.*, 34 (2003) 339–353.
[34] K. Jbilou, and A.J. Riquet. Projection methods for large Lyapunov matrix equations, *Linear Algebra Appl.*, 415 (2006) 344–358.
[35] C.R. Johnson. Numerical determination of the field of values of a general complex matrix, *SIAM J. Numer. Anal.* 15 (1978) 595–602.
[36] F. Kangal, K. Meerbergen, E. Mengi, and W. Michiels. *A subspace method for large scale eigenvalue optimization*, *SIAM J. Matrix Anal. Applic.*, 39 (2018) 48–82.
[37] C. Kenney and A. Laub. Controllability and stability radii for companion form systems, *Math. Control Signals Syst.*, 1 (1988) 239–256.
[38] L. Knizhnerman, and V. Simoncini. Convergence analysis of the extended Krylov subspace method for the Lyapunov equation. *Numer. Math.*, 118:567–586, 2011.
[39] D. Kressner. Subspace methods for computing the pseudospectral abscissa and the stability radius, *SIAM J. Matrix Anal. Appl.*, 35 (2014) 292–313.
[40] D. Kressner, and C. Tobler. Low-rank tensor Krylov subspace methods for parametrized linear systems. *SIAM J. Matrix Anal. Appl.*, 32:1288–1316, 2011.
[41] A. L'Afflitto, W.M. Haddad, and Q. Hui. Optimal control for linear and nonlinear semistabilization, *J. Franklin Inst.*, 352 (2015) 851–881.
[42] Y. Lin, and V. Simoncini. Minimal residual methods for large scale Lyapunov equations, *Appl. Numer. Math.*, 72 (2013) 52–71.
[43] MathWorks. *MATLAB User's Guide*, Natick, MA, 2020.
[44] E. Mengi. *Eigopt: Software for Eigenvalue Optimization*, 2014; available at `http://home.ku.edu.tr/∼emengi/software/eigopt.html`.
[45] E. Mengi, E.A. Yildirim, and M. Kilic. Numerical optimization of eigenvalues of Hermitian matrix functions, *SIAM J. Matrix. Anal. Applic.*, 35 (2014) 699–724.
[46] R. Olfati-Saber, and R.M. Murray. Consensus problems in networks of agents with switching topology and time-delays, *IEEE Trans. Autom. Control*, 49 (2004) 1520–1533.
[47] T. Penzl. A cyclic low-rank smith method for large sparse Lyapunov equations, *SIAM J. Sci. Comput.,* 21 (2000) 1401–1418.
[48] Y. Saad. Chebyshev acceleration techniques for solving nonsymmetric eigenvalue problems, *Math. Comp.*, 42 (1984) 567–588.
[49] Y. Saad. Numerical solution of large Lyapunov equation, in M.A. Kaashoek, J.H. Van Schuppen and A.C. Ran, eds., Signal Processing, Scattering, Operator Theory, and Numerical Methods, *Proc. Int. Symp. MTNS-89*, Vol. III, pages 71–89, Birkhauser, Boston, 1990.
[50] S.D. Shank, V. Simoncini, and D.B. Szyld. Efficient low-rank solutions of Generalized Lyapunov equations, *Numer. Math.*, 134 (2016) 327–342.
[51] V. Simoncini. Analysis of the rational Krylov subspace projection method for large-scale algebraic Riccati equations, *SIAM J. Matrix Anal. Applic.*, 37 (2016) 1655–1674.
[52] V. Simoncini. Computational Methods for Linear Matrix Equations, *SIAM Rev.*, 58 (2016) 377–441.
[53] V. Simoncini, and D.B. Szyld. Recent Computational Developments in Krylov Subspace Methods for Linear Systems, *Numer. Linear Algebra Appl.*, 14 (2007) 1–59.
[54] C. Van Loan. How near is a stable matrix to an unstable matrix? *Contemp. Math.*, 47 (1985) 465–478.
[55] L. Zhang, H.-Y. Fan, and E.K.-W. Chu. Inheritance properties of Krylov subspace methods for continuous-

time algebraic Riccati equations, *J. Comp. Appl. Math.*, 371 (2020) 112685.
[56] L. Zhang, H.-Y. Fan, and E.K.-W. Chu, Krylov subspace methods for discrete-time algebraic Riccati equations, *Appl. Numer. Math.*, 152 (2020) 499–510.