National Center for
Theoretical Sciences
Mathematics Division, Taiwan

NCTS/Math Technical Report
2017-003

Liping Zhang
Hung-Yuan Fan
Eric King-wah Chu

Inheritance Properties of Projection Methods for Continuous-Time Algebraic Riccati Equations

Inheritance Properties of Projection Methods for Continuous-Time Algebraic Riccati Equations

Liping Zhang* Hung-Yuan Fan† Eric King-wah Chu‡

Abstract

We consider the numerical solution of large-scale continuous-time algebraic Riccati equations by projection methods using Krylov subspaces. More importantly, we show that the solvability of the projected algebraic Riccati equation does not have to be assumed but can be *inherited*. Illustrative numerical examples are presented.

Keywords. algebraic Riccati equation, Krylov subspace, large-scale problem, LQG optimal control, projection method

AMS subject classifications. 15A24, 65F30, 93C05

1 Introduction

In this paper we consider the large-scale continuous-time algebraic Riccati equations (CAREs), especially the application of projection methods in general and the *inheritance* of solvability conditions from a CARE by the corresponding projected equation in particular. This inheritance property is obviously important but has not been investigated previously.

1.1 Algebraic Riccati Equations

Consider the linear time-invariant (LTI) control system in continuous-time:

$$\dot{x}(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t),$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{l\times n}$ with $m, l \leq n$. The linear quadratic Gaussian (LQG) control minimizes the cost functional $J_c(x,u) \equiv \int_0^\infty \left[x(t)^\top Hx(t) + u(t)^\top Ru(t)\right] dt$, with $H \equiv C^\top C \geq 0$ and $R > 0$. Here, a symmetric matrix is positive (semi-)definite, denoted by $M > 0$ (≥ 0), when all its eigenvalues are positive (non-negative). Also, $M > N$ ($M \geq N$) iff $M - N > 0$ (≥ 0). The corresponding optimal control $u(t) = -R^{-1}B^\top Xx(t)$ can be expressed in terms of the unique positive semi-definite stabilizing solution X of the CARE [9, 11, 35, 41]:

$$\mathcal{C}(X) \equiv A^\top X + XA - XGX + H = 0, \quad G = BR^{-1}B^\top \geq 0. \tag{1}$$

*Department of Mathematics, Zhejiang University of Technology, Hangzhou 310023, China; zhanglp@zjut.edu.cn

†Department of Mathematics, National Taiwan Normal University, Taipei 116, Taiwan; hyfan@math.ntnu.edu.tw

‡School of Mathematics, 9 Rainforest Walk, Monash University, Victoria 3800, Australia; eric.chu@monash.edu (corresponding author)

(We have neglected the estimation of the state x from filtering processes in real-life applications.)

Analogously for the LTI control system in discrete-time, the corresponding LQG control requires the solution of the discrete-time algebraic Riccati equation (DARE) [9, 12, 35, 41].

1.2 Previous Work

The solution of AREs (including CAREs and DAREs) is an active area of research due to its importance in optimal control and filtering. Many in control theory and applied mathematics worked on the problem, contributed dozens of methods [9, 11, 12, 35, 41]. Classical approaches made use of canonical forms, determinants and polynomial manipulation and state-of-the-art ones work in a numerically stable manner; see [11, 12] and their references for more details. A favourite approach reformulates the AREs as eigenvalue problems [36] and has been implemented in MATLAB (as the commands `care` and `dare`). Another favourite is the Newton-Kleinman method [32]. On modern algorithms for AREs of moderate dimensions, consult [11, 12, 36].

For control problems from PDEs and the balancing based model order reduction of large linear systems, large-scale CAREs, DAREs, Lyapunov and Stein equations have to be solved [2, 6, 7, 8, 23, 26, 27, 40, 47]. Solving the corresponding CARE may involve the invariant subspace of the Hamiltonian matrix

$$\mathcal{H} \equiv \begin{bmatrix} A & -G \\ -H & -A^\top \end{bmatrix}, \tag{2}$$

usually an prohibitively expensive exercise, as commented in [3].

Benner and his collaborators have done much on large-scale AREs [4, 5, 6, 7, 8, 47]. They built their methods on Newton's methods with ADI for the associated Lyapunov and Stein equations. (The initialization of Newton's method and the choice of parameters for the ADI are challenging.) Consult also [1, 26, 27] on various invariant or Krylov subspace methods. The structure-preserving doubling algorithm (SDA) [11, 12] has been adapted for large-scale problems [37], utilizing the structure in A and the low rank of H (i.e., $l \ll n$).

Inheritance Properties of Projection Methods

The CARE (1) is solvable and yields a unique stabilizing positive semi-definite solution X when the underlying control system is stabilizable and detectable [35, 41]. As in any numerical method, the solvability conditions should be reflected in the solution process. Apart from being mathematically elegant, this will avoid any unnecessary or ill-conditioned computation when the problem is (nearly) unsolvable. For example, in [36] or the command `care` in MATLAB, when the eigenvalues of $\mathcal{H}$ are on or near the imaginary axis, the computation should be carried out more carefully, or even abandoned, for the ill-conditioned or unsolvable problem.

The projection method is popular for large-scale CAREs, projects the original equation onto some subspace and produces a small projected CARE (pCARE). However, most papers have not elaborated on the solvability of the projected equation. One exception is [26], in which a deflation technique is applied to improve, hopefully, the pCARE. From [26], Jbilou assumed, for increasing d (the dimension of the Krylov subspace), that the projected systems "$\{H_d^\top, B_d\}$ is c-stabilizable and $\{H_d^\top, \widetilde{C}_1\}$ is c-detectable. These conditions ensure that the (projected) matrix equation has a unique symmetric positive semi-definite solution Y_d. If the preceding conditions are not satisfied we can use an implicit restart strategy to remove the unstable eigenvalues to obtain a c-stabilizable and c-detectable low-order model". However, it is unclear how the technique directly affects the solvability of the projected equation.

We quote from other representative literature on the *assumed* solvability of pCAREs.

(1) Heyouni and Jbilou [2008] [23, after (3.2)] "assume that the projected algebraic Riccati equation has a unique symmetric positive semi-definite and stabilizing solution".

(2) In Jbilou [2006] [27, after (2.5)], the pCARE was assumed to be c-stabilizable and c-detectable, thus uniquely solvable.

(3) In Lin and Simoncini [2015] [39, Section 1], the system matrix A is assumed to be stable. This limits the applicability of the results as one motive in the LQG control is stabilization.

(4) In Simoncini [2016] [48, after (2.2)], A is assumed to be passive (with a stable field of values) and dissipative (to pass on the stability to the projected system).

(5) In Simoncini, Szyld and Monsalve [2014] [49, Section 1], A is assumed to be stable or passive.

In Section 2.2, we investigate the inheritance of solvability conditions of projection methods for the CAREs, or when stabilizability, detectability and other conditions of the original control system are passed onto, or inherited by, the projected system.

1.3 Main Contributions

In this paper, we prove the unique solvability of the original ARE is *inherited* by the projected equation, under the condition that $\|\check{x}_1^\top r_k\|$ (in (13)) and $\|r_k \check{y}_2\|$ (in (18)) are respectively small relative to $\tau(A, B)$ and $\tau(A^\top, C^\top)$ (from (10b)), where r_k denotes the Arnoldi residual (in (5)). Other inheritance properties are presented in terms of the stability radius and the perturbation theory, with the latter independent of r_k. We have only made a start on the inheritance properties, leaving many questions unanswered, especially when $\|r_k\|$ does not diminish.

1.4 Organization of Paper

We consider projection methods for CAREs in Section 2, elaborating on the inheritance of some solvability conditions. Accuracy is considered in Section 3, some numerical examples are in Section 4 and we conclude in Section 5.

Notations. The 2- and F-norms are denoted by $\|\cdot\|$ and $\|\cdot\|_F$ respectively. Transposes and inverses are indicated respectively by $(\cdot)^\top$ and $(\cdot)^{-1}$, with the latter assuming invertibility implicitly. Lazily, we abbreviate $(A^\top)^{-1}$ to $A^{-\top}$. The identity matrix is I, occasionally with a subscript for its dimension. The maximum and minimum singular values are denoted respectively by $\sigma_{\max}(\cdot)$ and $\sigma_{\min}(\cdot)$, and the condition number $\kappa(A) \equiv \sigma_{\max}(A)/\sigma_{\min}(A)$. The spectrum is denoted by $\Lambda(\cdot)$ and the maximum eigenvalue by $\lambda_{\max}(\cdot)$. The field of real and complex numbers are respectively $\mathbb{R}$ and $\mathbb{C}$, with $\mathbb{C}_+$ being the closed right plane. The sets of real $m \times n$ matrices and real symmetric $n \times n$ matrices are respectively $\mathbb{R}^{m\times n}$ and $\mathbb{S}_n$. Occasionally, we abbreviate $A \leq C$ and $B \leq C$ as $A, B \leq C$.

2 Projection Methods

Projection methods are applicable when the solution X is numerically low-rank. For AREs, this is well known when H is low-rank with $l \ll n$ [3, 7, 8, 37]. There are many Krylov subspace methods, with $\mathcal{K}_k(\mathcal{M}, Z) \equiv \mathrm{span}\{Z, \mathcal{M}Z, \cdots, \mathcal{M}^{k-1}Z\}$ for $\mathcal{M} = A^{\pm\top}$ and $Z = C^\top$.

The following connection between SDA and $\mathcal{K}_{2^k}(A_\gamma^{-T}, A_\gamma^{-T}C^T)$ (with $A_\gamma \equiv A - \gamma I$ for $\gamma > 0$) has been suggested in [37]. Via the Cayley transform, the CARE in (1) is equivalent to a DARE. The SDA [11, 12, 37] can then be applied, corresponding to the Krylov subspace $\mathcal{K}_k(\mathcal{M}, Z)$ with

$$\mathcal{M} \equiv (A_\gamma^\top + HA_\gamma^{-1}G)^{-1}(A_{-\gamma}^\top + HA_\gamma^{-1}G), \quad Z \equiv A_\gamma^{-\top}C^\top.$$

For computation, we have

$$\begin{aligned} A_\gamma^{-\top}A_{-\gamma}^\top &= A_\gamma^{-\top}(A_\gamma^\top + 2\gamma I) = I + 2\gamma A_\gamma^{-\top}, \\ \mathcal{M} &= I + 2\gamma(A_\gamma^\top + HA_\gamma^{-1}G)^{-1} = I + 2\gamma A_\gamma^{-\top} - 2\gamma A_\gamma^{-\top}H(A_\gamma + GA_\gamma^{-\top}H)^{-1}GA_\gamma^{-\top}, \end{aligned}$$

by the Sherman-Morrison-Woodbury formula. Consequently, we have

$$\mathcal{K}_k(\mathcal{M}, A_\gamma^{-\top}C^\top) \subseteq \mathcal{K}_k(A_\gamma^{-\top}, A_\gamma^{-\top}C^\top) \subseteq \mathcal{K}_{k+1}(A_\gamma^{-\top}, C^\top). \tag{3}$$

Our numerical experience shows that the accuracy of approximate solutions and efficiency are similar with $\mathcal{K}_k(A_\gamma^{-\top}, A_\gamma^{-\top}C^\top)$ and $\mathcal{K}_k(A_\gamma^{-\top}, C^\top)$. For the rest of the paper, we shall choose the latter, which is simpler to analyze and apply.

From the Arnoldi process with $V_0 \equiv C^\top$ (orthonormalized), we have the Arnoldi relationship:

$$A_\gamma^{-\top}V_k = V_k\Omega_k + \widetilde{v}_{k+1}\widetilde{r}_k^\top, \quad V_{k+1} = [V_k, \widetilde{v}_{k+1}], \tag{4}$$

where Ω_k is upper block Hessenberg and $V_{k+1}^\top V_{k+1} = I$. After rearrangement and manipulation [13, 39, 48], (4) leads to the Arnoldi relationship for $A^\top$:

$$\begin{aligned} & (A-\gamma I)^{-\top}V_k = V_k\Omega_k + \widetilde{v}_{k+1}\widetilde{r}_k^\top \Leftrightarrow V_k = (A^\top - \gamma I)V_k\Omega_k + (A^\top - \gamma I)\widetilde{v}_{k+1}\widetilde{r}_k^\top \\ \Leftrightarrow\ & A^\top V_k\Omega_k = V_k(I + \gamma\Omega_k) - A_\gamma^\top\widetilde{v}_{k+1}\widetilde{r}_k^\top \Leftrightarrow A^\top V_k = V_k(I+\gamma\Omega_k)\Omega_k^{-1} - A_\gamma^\top\widetilde{v}_{k+1}\widetilde{r}_k^\top\Omega_k^{-1} \\ \Leftrightarrow\ & A^\top V_k = V_k\left(I + \gamma\Omega_k - V_k^\top A_\gamma^\top\widetilde{v}_{k+1}\widetilde{r}_k^\top\right)\Omega_k^{-1} - (I - V_kV_k^\top)A_\gamma^\top\widetilde{v}_{k+1}\widetilde{r}_k^\top\Omega_k^{-1} \\ \Leftrightarrow\ & A^\top V_k = V_k\Phi_k^\top + v_{k+1}r_k^\top, \end{aligned} \tag{5}$$

by the QR decomposition [18] $-(I - V_kV_k^\top)A_s^\top\widetilde{v}_{k+1} = v_{k+1}\beta$ with $v_{k+1}^\top v_{k+1} = I$, β full-rank, $\Phi_k^\top \equiv \left(I + \gamma\Omega_k - V_k^\top A_\gamma^\top\widetilde{v}_{k+1}\widetilde{r}_k^\top\right)\Omega_k^{-1}$ and $r_k^\top \equiv \beta\widetilde{r}_k^\top\Omega_k^{-1}$. Consequently, $V_{k+1} \equiv [V_k, v_{k+1}]$ retains the orthonormal columns and $\Phi_k = V_k^\top AV_k$. We refer to r_k as the Arnoldi residual which may be used to control the Arnoldi process (4). Note that Ω_k is block Hessenberg but not Φ_k.

Assuming the low-rank approximation $X_k \equiv V_kY_kV_k^\top \approx X$ with $Y_k^\top = Y_k$, $G_{11} \equiv V_k^\top GV_k$ and $H_{11} \equiv V_k^\top HV_k$, the Galerkin condition $V_k^\top\mathcal{C}(X_k)V_k = 0$ leads to the pCARE:

$$\widehat{C}(Y_k) \equiv \Phi_k^\top Y_k + Y_k\Phi_k - Y_kG_{11}Y_k + H_{11} = 0. \tag{6}$$

As in most publications on Krylov subspace methods, the Arnoldi residual r_k, which may persist in norm and does not diminish with respect to k, plays an important part in the convergence and accuracy of the algorithms. Consult also the discussions in Sections 2.1.1 and 5.1.

Remark 2.1 (Krylov Subspaces) *We base our analysis on the Arnoldi relationship in* (5) *thus the corresponding Krylov subspace* $\mathcal{K}_k(A^\top, C^\top)$ *may be appropriate for the projection method for CAREs, in addition to* $\mathcal{K}_k(A_\gamma^{-\top}, C^\top)$ *in* (3)*. There are other possibilities from [23, 26, 27, 39, 48, 49]. Comparing all the Krylov subspaces is a big project not within the scope of our paper. Our numerical experience suggests that the optimal choice of* $\mathcal{K}_k$ *depends on individual applications. Between* $\mathcal{K}_k(A^\top, C^\top)$ *and* $\mathcal{K}_k(A_\gamma^{-\top}, C^\top)$ *for the examples in Section 4, the execution times required are similar but the latter sometimes requires a lower-rank approximation. However, generating the subspaces with* $A_\gamma^{-\top}$ *is more expensive than* $A^\top$ *but the difference is minimal.*

Remark 2.2 (Efficiency) *With a sparse* A *containing a small number of nonzero elements on each column and* $V_k \in \mathbb{R}^{n\times d}$*, solving the CARE* (1) *via the Arnoldi process* (5) *for* $\mathcal{M} = A^\top$ *and the pCARE* (6) *is very efficient, involving* $O(n)+O(d^3)$ *flops. This operation count is competitive against other methods, when* X *is numerically low-rank and* d *is small. For the Krylov subspaces with* $\mathcal{M} = A_\gamma^{-\top}$*, appropriate structures in* A *or an efficient solver for* $A_\gamma^\top x = b$ *are required.*

2.1 Alternative Interpretation of Projection Methods

Let the orthogonal $P \equiv [P_1, P_2]$ with $P_1 \equiv V_k$. The solution X of the CARE (1) has the form

$$X = P \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{12}^\top & Y_{22} \end{bmatrix} P^\top = \sum_{i,j=1}^{2} P_i Y_{ij} P_j^\top,$$

where $Y_{21} = Y_{12}^\top$ and $Y_{ii}^\top = Y_{ii}$ $(i = 1, 2)$ from the symmetry of X, and $Y_{ij} = P_i^\top X P_j$ $(i, j = 1, 2)$. Obviously, Y_k in (6) is an approximation to Y_{11}. From the Arnoldi relationship (5), $A_{21} \equiv P_2^\top A P_1$ and $A_{22} \equiv P_2^\top A P_2$, we have

$$P^\top A P = \begin{bmatrix} P_1^\top A P_1 & P_1^\top A P_2 \\ P_2^\top A P_1 & P_2^\top A P_2 \end{bmatrix} = \begin{bmatrix} \Phi_k & r_k v_{k+1}^\top P_2 \\ A_{21} & A_{22} \end{bmatrix}. \tag{7}$$

Equation (7) is important — with $\|r_k\|$ relatively small, A is essentially triangularized and decoupled into two smaller subsystems represented by Φ_k and A_{22}. In fact, P_1 spans approximately an invariant subspace of A. Furthermore, the decomposition on the right of (7) is almost the observability canonical form for A, with Φ_k and A_{22} representing approximately the observable and unobservable subsystems, respectively. If $\{A, C\}$ is detectable, A_{22} is not far from being stable. Ignoring the "small" perturbation $r_k v_{k+1}^\top P_2$, the spectrum $\Lambda(A)$ equals to the disjoint union of $\Lambda(\Phi_k)$ and $\Lambda(A_{22})$. Thus, invertibility of A is passed onto Φ_k. We shall show later that stabilizability, detectability and other properties are similarly passed on.

A more precise statement on the relationship between the spectra of A and Φ_k requires the following theorem on the perturbation of eigenvalues [50]:

Theorem 2.1 (Theorem 2.10 in [50]) *Let* $P^\top AP$ *be partitioned as in* (7)*. If* $\delta = \mathrm{sep}(\Phi_k, A_{22}) \equiv \inf_{\|X\|=1} \|A_{22}X - X\Phi_k\| > 0$ *and* $\|A_{21}\|\|r_k\| < \delta^2/4$*, then there exists a unique* S *such that* $W = P_1 + P_2 S$ *spans an invariant subspace of* A*, and* $\Phi_k + r_k v_{k+1}^\top P_2 S$ *and* $A_{22} - S r_k v_{k+1}^\top P_2$ *are respectively the representations of* A *in* $\mathrm{span}(W)$ *and its orthogonal complement. The matrix* S *solves the Riccati equation* $A_{22}S - S\Phi_k - S r_k v_{k+1}^\top P_2 S + A_{21} = 0$ *with* $\|S\| < 2\delta^{-1}\|A_{21}\|$*.*

Remark 2.3 *From Theorem 2.1, $\Lambda(A)$ equals the disjoint union of $\Lambda(\Phi_k + r_k v_{k+1}^\top P_2 S)$ and $\Lambda(A_{22} - S r_k v_{k+1}^\top P_2)$. When A is nonsingular, for example, so is $\Phi_k + r_k v_{k+1}^\top P_2 S$ and if the nearby matrix Φ_k (after a perturbation of $\Delta \equiv -r_k v_{k+1}^\top P_2 S$ of magnitude $\|\Delta\| \leq \|r_k\|\|S\| < 2\delta^{-1}\|A_{21}\|\|r_k\|$) stays nonsingular, the invertibility of A is then inherited by Φ_k. This will be the case when $\|\Delta\|$ is sufficiently small, even if $\|r_k\|$ is large.*

With the Krylov subspace $\mathcal{K}_k(A_\gamma^{-\top}, C^\top)$, we have $\mathrm{span}\,(C^\top) \subseteq \mathrm{span}\,(P_1)$. From (7), together with $G_{ij} \equiv P_i^\top G P_j$ and $H_{ij} \equiv P_i^\top H P_j$ ($i, j = 1, 2$; all vanish except H_{11}), we have

$$
\begin{aligned}
& \mathcal{C}(X) = PP^\top \mathcal{C}(X) PP^\top \\
= \; & P \begin{bmatrix} \Phi_k^\top Y_{11} + Y_{11}\Phi_k - Y_{11} G_{11} Y_{11} - g_{11} + H_{11} + \widehat{H}_{11} & \widetilde{\Phi}_k Y_{12} + Y_{12}\widetilde{A}_{22} - g_{12} + \widehat{H}_{12} \\ * & \check{C}(Y_{22}) \end{bmatrix} P^\top \\
= \; & P \begin{bmatrix} \widehat{C}(Y_{11}) - g_{11} + \widehat{H}_{11} & \widetilde{\Phi}_k Y_{12} + Y_{12}\widetilde{A}_{22} - g_{12} + \widehat{H}_{12} \\ * & \check{C}(Y_{22}) \end{bmatrix} P^\top
\end{aligned} \tag{8}
$$

with $*$ representing the part of the symmetric matrix not needed to be specified and

$$
\begin{aligned}
\check{C}(Y_{22}) &\equiv A_{22}^\top Y_{22} + Y_{22} A_{22} - Y_{22} G_{22} Y_{22} - g_{22} + \widehat{H}_{22}, \\
\widetilde{\Phi}_k &\equiv \Phi_k - G_{11} Y_{11}, \quad \widetilde{A}_{22} \equiv A_{22} - G_{22} Y_{22}; \\
\widehat{H}_{11} &\equiv A_{21}^\top Y_{12}^\top + Y_{12} A_{21}, \quad g_{11} \equiv Y_{11} G_{12} Y_{12}^\top + Y_{12} G_{21} Y_{11} + Y_{12} G_{22} Y_{12}^\top; \\
\widehat{H}_{12} &\equiv Y_{11} r_k v_{k+1}^\top P_2 + A_{21}^\top Y_{22}, \quad g_{12} \equiv Y_{11} G_{12} Y_{22} + Y_{12} G_{21} Y_{12}; \\
\widehat{H}_{22} &\equiv P_2^\top v_{k+1} r_k^\top Y_{12} + Y_{12}^\top r_k v_{k+1}^\top P_2, \quad g_{22} \equiv Y_{12}^\top G_{11} Y_{12} + Y_{12}^\top G_{12} Y_{22} + Y_{22} G_{21} Y_{12}.
\end{aligned}
$$

If $Y_{12}, Y_{22} = 0$ then $Y_{11} = Y_k$. In general, from (8), Y_k is a good approximation to Y_{11} for an accurate projection method. This suggests an iterative method for X or Y_{ij} from the zero starting point, yielding Y_k in the first iteration with $Y_{12} = O(\|Y_{11} r_k\|)$ and $Y_{22} = O(\|Y_{12}^\top r_k\|)$ (which can be proved by the Newton-Kantorovich theorem, with details neglected).

2.1.1 When $\|r_k\|$ Is Not Small

Often the projection method produces an accurate approximation to the solution X of the CARE (1) when $\|r_k\|$ is not small. In fact, a small $\|r_k\|$ reflects the accurate approximation of an invariant subspace of A by $\mathrm{span}\,(P_1)$, which is more useful in solving linear systems or eigenvalue problems associated with A. For the solution of CAREs, we require X to be numerically low-rank and $\|P_2^\top X\| \leq \varepsilon_k$ for an appropriate Krylov subspace $\mathrm{span}\,(P_1)$, for the projection method to be applicable. We then observe that

$$\|Y_{12}\|, \|Y_{22}\| \leq \|[Y_{12},\, Y_{22}]\| = \|P_2^\top X\| \leq \varepsilon_k.$$

Together with (8), we obtain $\|\widehat{H}_{22}\|, \|Y_{11} r_k\| = O(\varepsilon_k)$, even though $\|r_k\|$ is not small. (Similarly, from (28), we see that the residual R_k of the approximate solution $X_k \equiv V_k Y_k V_k^\top$ from a projection method equals $\|Y_k r_k\| \approx \|Y_{11} r_k\|$.) From the unique solvability of (1), we deduce that Y_{11} solves (6) approximately but no conclusion can be drawn on the (exact) solvability of the pCARE or whether Y_k is stabilizing or definite (c.f. Theorem 2.5). Note that stabilizability and detectability are only sufficient for the solvability of CAREs, complicating any investigation. For example in

[26], Jbilou deflated the Krylov subspace so as to eliminate the unstable eigenvalues in Φ_k. This may be wasteful to eliminate unstable but controllable poles unnecessarily. Further discussions can be found in Sections 3.1 and 5.1.

Note that for the inheritance properties in Theorems 2.2–2.4, we require $\|\check{x}_1^\top r_k\|$ and $\|r_k \check{y}_2\|$ (potentially much smaller than $\|r_k\|$ with $\|\check{x}_1\|, \|\check{y}_2\| \le 1$) to be relatively small.

2.2 Inheritance of Solvability

Consider the projection method with $\mathcal{K}_k(A_\gamma^{-\top}, C^\top)$. Recall the Arnoldi relationship (5) or (7) and denote $B_i \equiv P_i^\top B$ and $C_i \equiv CP_i$ $(i = 1, 2)$. We have several approaches to investigate the solvability conditions of the CARE (1) and its projection (6).

2.2.1 Stabilizability

For all $s \in \mathbb{C}_+$ (the closed right plane), stabilizability of $\{A, B\}$ [24, 30, 41, 44] is equivalent to the full-rank (f.r.) conditions:

$$[sI - A,\, B] \text{ f.r.} \Leftrightarrow P^\top [sI - A,\, B] \begin{bmatrix} P & 0 \\ 0 & I_m \end{bmatrix} \text{ f.r.} \Leftrightarrow \left[\begin{array}{cc|c} sI - \Phi_k & -r_k v_{k+1}^\top P_2 & B_1 \\ -A_{21} & sI - A_{22} & B_2 \end{array}\right] \text{ f.r.}. \tag{9}$$

If the $r_k v_{k+1}^\top P_2$ term is considered to be an "error" or a "noise", then the stabilizability of $\{A,\, B\}$ is obviously inherited by the projected system $\{\Phi_k, B_1\}$. We expand the statement as follows:

(i) If $\{A,\, B\}$ is stabilizable but $\{\Phi_k,\, B_1\}$ is not, or equivalently $[sI - \Phi_k,\, B_1]$ is not full-rank for all $s \in \mathbb{C}_+$, then $\{A,\, B\}$ is within an $O(\|r_k\|)$ distance from unstabilizability. (We obtain the better result later with $\|r_k\|$ replaced by the potentially smaller $\|\check{x}_1^\top r_k\|$ in (13).)

(ii) With $\sigma_{\min}$ denoting the smallest singular value, we modify the definition of the distance of the system $\{A,\, B\}$ from uncontrollability in [19, 44] to that for unstabilizability:

$$\begin{aligned} \tau(A, B) &\equiv \min\{\|[\delta A, \delta B]\|_F : \{A + \delta A, B + \delta B\} \text{ is not stabilizable}\} && \text{(10a)} \\ &= \min_{s \in \mathbb{C}_+} \sigma_{\min}(A - sI,\, B). && \text{(10b)} \end{aligned}$$

Apply the properties of singular values [18] to the matrix on the right of (9), we have

$$\begin{aligned} \tau(A, B) &= \min_{s \in \mathbb{C}_+} \sigma_{\min} \left[\begin{array}{cc|c} sI - \Phi_k & B_1 & -r_k v_{k+1}^\top P_2 \\ \hline -A_{21} & B_2 & sI - A_{22} \end{array}\right] \\ &= \min_{s \in \mathbb{C}_+} \min_{\|[x_1^\top, x_2^\top]\| = 1} \left\| [x_1^\top, x_2^\top] \left[\begin{array}{cc|c} sI - \Phi_k & B_1 & -r_k v_{k+1}^\top P_2 \\ \hline -A_{21} & B_2 & sI - A_{22} \end{array}\right] \right\| \\ &\le \min_{s \in \mathbb{C}_+} \min_{\|[x_1^\top, x_2^\top]\| = 1} \left\{ \left\| [x_1^\top, x_2^\top] \left[\begin{array}{cc|c} sI - \Phi_k & B_1 & 0 \\ \hline -A_{21} & B_2 & sI - A_{22} \end{array}\right] \right\| + \left\| x_1^\top r_k \right\| \right\}. && (11) \end{aligned}$$

Optimizing the first term in (11) with $x \equiv [x_1^\top, x_2^\top]^\top = [\check{x}_1^\top, \check{x}_2^\top]^\top$ leads to

$$\begin{aligned}
\tau(A,B) &\leq \min_{s\in\mathbb{C}_+} \min_{\|x\|=1} \left\| x^\top \left[\begin{array}{cc|c} sI-\Phi_k & B_1 & 0 \\ \hline -A_{21} & B_2 & sI-A_{22} \end{array}\right] \right\| + \|\check{x}_1^\top r_k\| \\
&\leq \min_{s\in\mathbb{C}_+} \min_{\|x_1\|=1} \left\| [x_1^\top, 0] \left[\begin{array}{cc|c} sI-\Phi_k & B_1 & 0 \\ \hline -A_{21} & B_2 & sI-A_{22} \end{array}\right] \right\| + \|\check{x}_1^\top r_k\| \\
&= \tau(\Phi_k, B_1) + \|\check{x}_1^\top r_k\|. \qquad (12)
\end{aligned}$$

In summary, we have the following theorem:

Theorem 2.2 (Inheritance of Stabilizability) *With $\tau(A,B)$ measuring the distance from unstabilizability as defined in* (10) *and $\check{x}_1$ as in (12), assuming $\tau(A,B) > \|\check{x}_1^\top r_k\|$, we have*

$$\tau(\Phi_k, B_1) \geq \tau(A,B) - \|\check{x}_1^\top r_k\| > 0. \qquad (13)$$

Hence $\{\Phi_k, B_1\}$ inherits the stabilizability of $\{A,B\}$, after excluding the possibility when $\tau(A,B)$ is relatively small or $\{A,B\}$ within a distance of $\|\check{x}_1^\top r_k\|$ from being unstabilizable.

In other words, when $\|\check{x}_1^\top r_k\|$ is small enough, the stabilizability of $\{A,B\}$ is inherited by the projected system $\{\Phi_k, B_1\}$.

2.2.2 Detectability

With a sufficiently small $\|r_k\|$, $P^\top AP$ is essentially lower triangular in (7). Consequently from $\operatorname{span} C^\top \subseteq \mathcal{K}_k(A_\gamma^{-\top}, C^\top)$ and (7), for all $s \in \mathbb{C}_+$, detectability of $\{A,C\}$ is equivalent to

$$\begin{aligned}
\begin{bmatrix} sI-A \\ C \end{bmatrix} \text{ f.r.} \;&\Leftrightarrow\; \begin{bmatrix} P^\top & 0 \\ 0 & I_l \end{bmatrix} \begin{bmatrix} sI-A \\ C \end{bmatrix} P \text{ f.r.} \Leftrightarrow \left[\begin{array}{cc} sI-\Phi_k & -r_k v_{k+1}^\top P_2 \\ -P_2^\top AP_1 & sI-P_2^\top AP_2 \\ \hline CP_1 & CP_2 \end{array}\right] \text{ f.r.} \\
&\Leftrightarrow \left[\begin{array}{cc} sI-\Phi_k & 0 \\ -A_{21} & sI-A_{22} \\ \hline C_1 & 0 \end{array}\right] \text{ f.r.} \Leftrightarrow \left[\begin{array}{cc} sI-\Phi_k & 0 \\ 0 & sI-A_{22} \\ \hline C_1 & 0 \end{array}\right] \text{ f.r.}, \qquad (14)
\end{aligned}$$

implying $\{\Phi_k^\top, C_1^\top\}$ is detectable, the corresponding inheritance property and the stability of A_{22}.

More rigorously, when considering the minimum singular value similar to (14), we have

$$\begin{aligned}
\tau(A^\top, C^\top) &\equiv \min_{s\in\mathbb{C}_+} \sigma_{\min} \begin{bmatrix} sI-A \\ C \end{bmatrix} = \min_{s\in\mathbb{C}_+} \sigma_{\min} \left[\begin{array}{cc} sI-\Phi_k & -r_k v_{k+1}^\top P_2 \\ -A_{21} & sI-A_{22} \\ \hline CP_1 & 0 \end{array}\right] \\
&= \min_{s\in\mathbb{C}_+} \min_{\|[y_1^\top, y_2^\top]\|=1} \left\| \left[\begin{array}{cc} sI-\Phi_k & -r_k v_{k+1}^\top P_2 \\ -A_{21} & sI-A_{22} \\ \hline CP_1 & 0 \end{array}\right] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\| \\
&\leq \min_{s\in\mathbb{C}_+} \min_{\|[y_1^\top, y_2^\top]\|=1} \left\| \left[\begin{array}{cc} sI-\Phi_k & 0 \\ -A_{21} & sI-A_{22} \\ \hline CP_1 & 0 \end{array}\right] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\| + \|r_k v_{k+1}^\top P_2 \widetilde{y}_2\|, \qquad (15)
\end{aligned}$$

where $y \equiv [y_1^\top, y_2^\top]^\top = [\widetilde{y}_1^\top, \widetilde{y}_2^\top]^\top$ optimizes the first term in (15).

With $\check{y}_2 \equiv v_{k+1}^\top P_2 \widetilde{y}_2$ and assuming

$$\tau(A^\top, C^\top) > \|r_k \check{y}_2\|, \tag{16}$$

(15) implies that A_{22} is stable. Denote $\eta(s) \equiv -(sI - A_{22})^{-1}A_{21}$, $\Psi(s) \equiv \begin{bmatrix} I & 0 \\ \eta(s) & I \end{bmatrix}$ and $z \equiv \Psi(s)y$, we obtain

$$\begin{aligned}
\tau(A^\top, C^\top) \quad &\leq \quad \min_{s\in\mathbb{C}_+} \min_{y\neq 0} \left\{ \left\| \left[\begin{array}{cc} sI - \Phi_k & 0 \\ 0 & sI - A_{22} \\ \hline C_1 & 0 \end{array} \right] \Psi(s) y \right\| \cdot \frac{1}{\|y\|} \right\} + \|r_k \check{y}_2\| \\
&= \quad \min_{s\in\mathbb{C}_+} \min_{y\neq 0} \left\{ \left\| \left[\begin{array}{cc} sI - \Phi_k & 0 \\ 0 & sI - A_{22} \\ \hline C_1 & 0 \end{array} \right] z \right\| \cdot \frac{\|\Psi(s)\|}{\|\Psi(s)\|\|y\|} \right\} + \|r_k \check{y}_2\|
\end{aligned}$$

From $\|\Psi(s)\|\|y\| \geq \|\Psi(s)y\| = \|z\|$ and the positive constant

$$\eta^* \equiv \max_{s\in\mathbb{C}_+} \|\Psi(s)\| = \max_{s\in\mathbb{C}_+} \left\| \begin{bmatrix} I & 0 \\ -(sI - A_{22})^{-1}A_{21} & I \end{bmatrix} \right\| \leq 1 + \max_{s\in\mathbb{C}_+} \|(sI - A_{22})^{-1}A_{21}\|, \tag{17}$$

we have

$$\begin{aligned}
\tau(A^\top,\, C^\top) \quad &\leq \quad \eta^* \cdot \min_{s\in\mathbb{C}_+} \min_{y\neq 0} \left\{ \left\| \left[\begin{array}{cc} sI - \Phi_k & 0 \\ 0 & sI - A_{22} \\ \hline C_1 & 0 \end{array} \right] z \right\| \cdot \frac{1}{\|z\|} \right\} + \|r_k \check{y}_2\| \\
&= \quad \eta^* \cdot \min_{s\in\mathbb{C}_+} \sigma_{\min} \left[\begin{array}{cc} sI - \Phi_k & 0 \\ 0 & sI - A_{22} \\ \hline C_1 & 0 \end{array} \right] + \|r_k \check{y}_2\| \\
&\leq \quad \eta^* \cdot \min_{s\in\mathbb{C}_+} \sigma_{\min} \begin{bmatrix} sI - \Phi_k \\ C_1 \end{bmatrix} + \|r_k \check{y}_2\| \\
&\leq \quad \eta^* \cdot \tau(\Phi_k^\top,\, C_1^\top) + \|r_k \check{y}_2\|.
\end{aligned} \tag{18}$$

In summary, we have the following theorem:

Theorem 2.3 (Inheritance of Detectability) *With $\tau(A^\top, C^\top)$ measuring the distance from undetectability as in* (10)*, $\check{y}_2$ from (16) and $\eta^* > 0$ from (17), we have*

$$\eta^* \cdot \tau(\Phi_k^\top, C_1^\top) \geq \tau(A^\top, C^\top) - \|r_k \check{y}_2\|.$$

Hence, $\{\Phi_k^\top, C_1^\top\}$ inherits the detectability of $\{A^\top, C^\top\}$ if (16) *holds, excluding the possibility of $\tau(A^\top, C^\top)$ being relatively small or $\{A^\top, C^\top\}$ within a distance of $\|r_k \check{y}_2\|$ from undetectability.*

In other words, when the CARE (1) is uniquely solvable under the assumption of stabilizability and detectability, the pCARE (6) inherits the unique solvability, when $\{A, B, C\}$ is not within an $\|\check{x}_1^\top r_k\|$ or $\|r_k \check{y}_2\|$ distance from unstabilizability or undetectability, respectively. Note that $\|\check{x}_1^\top r_k\|$ and $\|r_k \check{y}_2\|$ (with $\|\check{x}_1\|, \|\check{y}_2\| \leq 1$) may potentially be much smaller than $\|r_k\|$. The inheritance of stabilizability and detectability for other projection methods with alternative Krylov subspaces has also been proved, when (5) holds with V_k, r_k and v_{k+1} defined differently.

2.2.3 Inheritance of Solvability from Hamiltonian Formulation

Stabilizability and detectability are only sufficient for the existence of a unique stabilizing positive semi-definite solution to the CARE (1) [35, 41]. They guarantee the separation of the eigenvalues of the Hamiltonian matrix $\mathcal{H}$ in (2), or the nonexistence of eigenvalues on the imaginary axis. Considering the solution of CAREs directly in terms of $\mathcal{H}$, we shall investigate the effect of projection on the stability of $\mathcal{H}$, or the distance of its spectrum from the imaginary axis.

To analyze stability, one good tool is the *stability radius* or *margin* [10, 22, 24, 28, 30, 45, 52]:

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

It is well known that $\Lambda(\mathcal{H})$ is the union of the the stable and antistable subspectra $\Lambda(A-GX)$ and $\Lambda(-(A-GX))$ respectively [35, 41], so the stability radius of $A-GX$:

$$\psi(A-GX) = \min_{\omega\in\mathbb{R}}\sigma_{\min}(A-GX-i\omega I),$$

the magnitude of the perturbation to $A-GX$ which will push it to instability, is a good measure of solvability of the CARE (1). The analogous quantity for the pCARE (6) is

$$\psi(\Phi_k - G_{11}Y_k) \equiv \min_{\omega\in\mathbb{R}}\sigma_{\min}(\Phi_k - G_{11}Y_k - i\omega I).$$

Let $R_k \equiv \mathcal{C}(X_k)$, $\mathcal{L}(\cdot) \equiv A_c^\top(\cdot) + (\cdot)A_c$ with $A_c \equiv A-GX$ and $c_1 \equiv 2\|\mathcal{L}^{-1}\|\|G\|$. From [29, Theorem 2] (or [48, Theorem 4.1]), when

$$\|X_k - X\| < (3\|G\|\|\mathcal{L}^{-1}\|)^{-1}, \quad 4\|G\|\|\mathcal{L}^{-1}\|^2\|R_k\| < 1, \tag{19}$$

we have

$$\|X_k - X\| \le 2\|\mathcal{L}^{-1}\|\|R_k\|. \tag{20}$$

With $z = [z_1^\top, z_2^\top]^\top = [\check{z}_1^\top, \check{z}_2^\top]^\top$ optimizing the first term in (21), techniques similar to those in Section 2.2.1 and 2.2.2 produce

$$\begin{aligned}
\psi(A-GX) &\le \psi(A-GX_k) + \|G(X_k - X)\| \le \psi(A-GX_k) + c_1\|R_k\| \\
&= \psi\begin{bmatrix} \Phi_k - G_{11}Y_k & r_k v_{k+1}^\top P_2 \\ A_{21} - P_2^\top G P_1 Y_k & A_{22} \end{bmatrix} + c_1\|R_k\| \\
&\le \min_{\omega\in\mathbb{R}}\min_{\|z\|=1}\left\| z^\top \begin{bmatrix} \Phi_k - G_{11}Y_k - i\omega I & 0 \\ A_{21} - P_2^\top G P_1 Y_k & A_{22} - i\omega I \end{bmatrix}\right\| + \|\check{z}_1^\top r_k\| + c_1\|R_k\| \qquad (21)\\
&\le \min_{\omega\in\mathbb{R}}\min_{\|z_1\|=1}\left\| [z_1^\top, 0] \begin{bmatrix} \Phi_k - G_{11}Y_k - i\omega I & 0 \\ A_{21} - P_2^\top G P_1 Y_k & A_{22} - i\omega I \end{bmatrix}\right\| + \|\check{z}_1^\top r_k\| + c_1\|R_k\| \\
&= \min_{\omega\in\mathbb{R}}\min_{\|z_1\|=1}\left\| z_1^\top(\Phi_k - G_{11}Y_k - i\omega I)\right\| + \|\check{z}_1^\top r_k\| + c_1\|R_k\| \\
&= \psi(\Phi_k - G_{11}Y_k) + \|\check{z}_1^\top r_k\| + c_1\|R_k\|.
\end{aligned}$$

The following theorem summarizes the inheritance property associated with ψ.

Theorem 2.4 (Inheritance of Stability Radius) *Assume* (19) *and that a unique solution Y_k exists for the pCARE* (6)*. With the stability radius ψ, we have*

$$\psi(\Phi_k - G_{11}Y_k) > \psi(A-GX) - \|\check{z}_1^\top r_k\| - c_1\|R_k\|,$$

where $\check{z}_1$ is defined in (21)*. In other words, $\Phi_k - G_{11}Y_k$ inherits the stability of $A - GX$ if $\psi(A - GX) > \|\check{z}_1^\top r_k\| + c_1\|R_k\|$, excluding the possibility of $\psi(A - GX)$ being too small or $A - GX$ within a distance of $\|\check{z}_1^\top r_k\| + c_1\|R_k\|$ from instability.*

2.2.4 Inheritance of Solvability from Perturbation Theory

We consider the solvability of the pCARE (6), equivalent to

$$A^\top X_k + X_k A - X_k G X_k + (H - R_k) = 0 \tag{22}$$

when $X_k = P_1 Y_k P_1^\top$. This inheritance property comes from the perturbation theory of CAREs, with (22) borrowing the solvability of the neighbouring (1). Note that the theorem contains a computable check for the solvability of the pCARE.

Theorem 2.5 *Let X_k from a projection method approximate the unique stabilizing solution X to CARE* (1) *with the residual $R_k = \mathcal{C}(X_k)$. Define the Lyapunov operator $\mathcal{L}(\cdot)$ as in* (19)*. With $X_k = P_1 Y_k P_1^\top$, Y_k is the unique solution to the pCARE (6) if*

$$4\|\mathcal{L}^{-1}\|\|\mathcal{L}^{-1}R_k\|\|G\| < 1. \tag{23}$$

Furthermore, the error satisfies

$$\|X_k - X\| \leq \frac{2\|\mathcal{L}^{-1}R_k\|}{1 + \sqrt{1 - 4\|\mathcal{L}^{-1}\|\|\mathcal{L}^{-1}R_k\|\|G\|}}.$$

Proof. Substituting $X_k = X + \delta X_k$ into the perturbed CARE (22) gives rise to

$$\mathcal{L}(\delta X_k) = \delta X_k G \delta X_k + R_k,$$

or equivalently,

$$\delta X_k = \mathcal{L}^{-1}(\delta X_k G \delta X_k) + \mathcal{L}^{-1}R_k \equiv \Upsilon(\delta X_k).$$

Then $\Upsilon : \mathcal{B}_\xi \to \mathcal{B}_\xi$ is a continuous mapping defined on the compact and convex set

$$\mathcal{B}_\xi = \{\delta X \in \mathbb{S}_n : \|\delta X\| \leq \xi, P_1^\top \delta X P_2 = -Y_{12}, P_2^\top \delta X P_2 = -Y_{22}\},$$

where

$$\xi = \frac{1 - \sqrt{1 - 4\|\mathcal{L}^{-1}\|\|\mathcal{L}^{-1}R_k\|\|G\|}}{2\|\mathcal{L}^{-1}\|\|G\|} = \frac{2\|\mathcal{L}^{-1}R_k\|}{1 + \sqrt{1 - 4\|\mathcal{L}^{-1}\|\|\mathcal{L}^{-1}R_k\|\|G\|}}$$

is positive if

$$4\|\mathcal{L}^{-1}\|\|\mathcal{L}^{-1}R_k\|\|G\| < 1. \tag{24}$$

Consequently, the mapping Υ has a fixed point $\delta X_k \in \mathcal{B}_\xi$ so that $X_k = P_1(Y_{11} + P_1^\top \delta X_k P_1)P_1^\top \equiv P_1 Y_k P_1^\top$ satisfies the perturbed CARE (22).

Moreover from [51, Lemma 2.1], the matrix $A - GX_k$ is c-stable if the condition (24) holds, hence $X_k = P_1 Y_k P_1^\top$ must be the unique stabilizing solution to the perturbed CARE (22). Since $P_1^\top R_k P_1 = 0$, it follows immediately that Y_k is the unique solution to the pCARE (6). □

Remark 2.4 *From the above proof and Theorem 2.4, if* $\psi(A - GX_k) > \|\check{z}_1^\top r_k\|$*, we obtain*

$$\psi(\Phi_k - G_{11}Y_k) \geq \psi(A - GX_k) - \|\check{z}_1^\top r_k\| > 0. \tag{25}$$

We then conclude that Y_k *is the unique stabilizing solution to the pCARE.*

The assumption in (25) *is not required if we are satisfied with the uniqueness of* Y_k *but not whether it is stabilizing. Importantly, the inheritance of solvability of the pCARE is dependent only on* R_k *in Theorem 2.5 but independent of* r_k*, unlike other inheritance properties.*

3 Accuracy of Projection Methods

There are some interesting results on the errors of projection methods in [48], such as the error in X_k (Theorem 4.1), the perturbed CARE which X_k satisfies (Proposition 5.1) and the perturbation of the corresponding stable Hamiltonian invariant subspace (Proposition 6.1). A link between projection methods and model order reduction can be found in [48, Section 3].

From the Arnoldi relationship (5) and $C^\top \in \operatorname{span} P_1$, we have

$$\begin{aligned} P_2^\top \mathcal{C}(X_k) &= P_2^\top (A^\top V_k Y_k V_k^\top + V_k Y_k V_k^\top A + V_k Y_k V_k^\top G V_k Y_k V_k^\top + H) \\ &= P_2^\top (V_k \Phi_k^\top + v_{k+1} r_k^\top) Y_k V_k^\top = P_2^\top v_{k+1} r_k^\top Y_k P_1^\top. \end{aligned} \tag{26}$$

Together with the Galerkin condition $P_1^\top \mathcal{C}(X_k) P_1 = 0$, the residual $R_k \equiv \mathcal{C}(X_k)$ satisfies

$$\begin{aligned} R_k &= PP^\top \mathcal{C}(V_k Y_k V_k^\top) PP^\top = P \begin{bmatrix} P_1^\top \mathcal{C}(X_k) P_1 & P_1^\top \mathcal{C}(X_k) P_2 \\ P_2^\top \mathcal{C}(X_k) P_1 & P_2^\top \mathcal{C}(X_k) P_2 \end{bmatrix} P^\top \\ &= P \begin{bmatrix} 0 & Y_k r_k v_{k+1}^\top P_2 \\ P_2^\top v_{k+1} r_k^\top Y_k & 0 \end{bmatrix} P^\top, \end{aligned} \tag{27}$$

thus the residual satisfies $R_k = v_{k+1} r_k^\top Y_k P_1^\top + P_1 Y_k r_k v_{k+1}^\top$, implying $\|R_k\| \leq 2\|Y_k r_k\|$. Note that (27) represents the same result as [48, Proposition 5.1]. When considering the singular values of the matrix on the right of (27), we have the slightly better result:

$$\|R_k\|^2 = \lambda_{\max}\left\{Y_k r_k r_k^\top Y_k\right\} = \|Y_k r_k\|^2. \tag{28}$$

This is interesting, especially when $\|r_k\|$ is large but $\|R_k\|$ is small. Notice that Y_k is the coefficient matrix in $X_k = V_k Y_k V_k^\top$ for the Krylov basis vectors in V_k for the projection method. Assume that the method is producing more accurate approximate solutions for increasing k, as the Krylov subspaces improve in accuracy by adding less significant components. This corresponds to $(Y_k)_{ij} \to 0$ as $i, j \to \infty$, thus $\kappa(Y_k) \to \infty$ as $k \to \infty$. It also indicates why $\|R_k\| = \|Y_k r_k\|$ can still be small even when r_k stagnates and is significant in the last few components. Note also the result $\|X_k - X\| \leq 2\|\mathcal{L}^{-1}\|\|R_k\|$ in (20) from [29, Theorem 2].

For a lower bound of $\|R_k\|$, from (27), we have

$$r_k^\top = v_{k+1}^\top R_k P_1 Y_k^{-1} \quad \Rightarrow \quad \|r_k\| = \|v_{k+1}^\top R_k P_1 Y_k^{-1}\| \leq \|Y_k^{-1}\| \cdot \|R_k\|, \tag{29}$$

leading to the equivalence relationship $\|R_k\|/\|Y_k\| \leq \|r_k\| \leq \|Y_k^{-1}\| \cdot \|R_k\|$. Equality results linking R_k and r_k already exist in (28) and (29). For the relative residual $\rho_k \equiv \|R_k\|/\|X_k\| = \|R_k\|/\|Y_k\|$, we also have

$$\rho_k \leq \|r_k\| \leq \kappa(Y_k)\rho_k, \quad \frac{\|r_k\|}{\kappa(Y_k)} \leq \rho_k \leq \|r_k\|. \tag{30}$$

The bounds involving Y_k^{-1} are impractical when Y_k is ill-conditioned or numerically low-rank.

From (26), we have the important fact that the CARE has a (hopefully small) $O(\|Y_k r_k\|)$ footprint on span P_2. Interestingly and naturally, the condition of Y_k is related to the residuals and r_k, as well as the solvability of the original and projected Riccati equations.

3.1 Problems with Arnoldi Process and Truncation

Projection methods may reduce the sizes of the AREs thus the associated workload, but have their fair shares of problems.

Breakdown An important phenomenon of the Arnoldi process is "breakdown", as the orthogonality of V_k is lost when contaminated by round-off errors. See, e.g., [17] for details. For the solution of linear systems and eigenvalue problems, breakdowns mean that some invariant subspace of A is well approximated, thus beneficial for the corresponding projection method. For the solution of CAREs, the Arnoldi residual r_k (in (4)) will be very small. This will be favourable for our investigation of the inheritance properties in Section 2.2.

Stagnation Contrary to breakdowns, "stagnation" generally means the Arnoldi process does not make any "progress". For linear systems, the solution is not approximated well until a very large Krylov subspace is generated, rendering the projection method unsuitable. Of course, it may be remedied by a more appropriate Krylov subspace. For CAREs, we have experienced having r_k persisted to be significant and not diminishing. However in some of these cases (as in Examples 1–3, Section 4.2), the projection method still produces accurate approximate solutions efficiently. In Section 2.2, we see that our analysis for the inheritance properties in Theorems 2.2–2.4 may fail (not for Example 2 from [26]) but the residual R_k in Theorem 2.5 can be small.

For the projection method to work for a small k, $\|R_k\| = \|Y_k r_k\|$ (from (28)) is small but Y_k is dominated by its upper left corner components. Stagnation means that r_k is dominated by the bottom components, forcing it to be large until k approaches n. These combine to explain why $\|R_k\| = \|Y_k r_k\|$ is small even when r_k is significant. Independent of r_k, if $\|R_k\|$ refuses to diminish, the solution X may not be numerically low-rank and no existing numerical methods may work for large-scale CAREs. Solving such a CARE may be too ambitious, as the large high-rank X cannot even be stored explicitly. We may instead estimate the feedback gain $F = -R^{-1}B^\top X \in \mathbb{R}^{m\times n}$, which is of lower dimensions than X.

We need to conduct further investigation of the inheritance properties when r_k stagnates, especially on how the solvability of the pCAREs can be guaranteed.

Reorthogonalization/Truncation For breakdowns or the loss of linear dependence of V_k in the Arnoldi process, reorthogonalization [17] can be applied in the corresponding Gram-Schmidt process. However in our experience, the QR factorization implemented by MATLAB produces satisfactory results without reorthogonalization.

For Section 4, the QR factorization with column pivoting (similar to the rank-revealing QR factorization) [18] on $[V_k,\, A_\gamma^{-\top} V_k]$ is applied and insignificant components in R (and the corresponding components in Q) are truncated, controlled by the truncation tolerance tol_def $= \epsilon > 0$. It is similar to deflation in [27]. Without truncation, the columns of V_k are nearly linearly dependent and we have the QR decomposition (ignoring some permutation matrix), for some small ϵ:

$$V_k = QR = [Q_1,\, Q_2] \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}, \quad \|R_{22}\| \leq \epsilon.$$

Retaining R_{22} and Q_2, the Hamiltonian matrix $\mathcal{H}$ in (2) for the pCARE (6) satisfies, with $\widetilde{\Phi}_k \equiv Q_1^\top A Q_1$, $\widetilde{G}_{11} \equiv Q_1^\top G Q_1$ and $\widetilde{H}_{11} \equiv Q_1^\top H Q_1$:

$$\begin{aligned}
\widetilde{\mathcal{H}} &\equiv \begin{bmatrix} R^\top & \\ & R^\top \end{bmatrix} \begin{bmatrix} Q^\top & \\ & Q^\top \end{bmatrix} \mathcal{H} \begin{bmatrix} Q & \\ & Q \end{bmatrix} \begin{bmatrix} R & \\ & R \end{bmatrix} \\
&= \begin{bmatrix} R^\top & \\ & R^\top \end{bmatrix} \begin{bmatrix} Q^\top A Q & -Q^\top G Q \\ -Q^\top H Q & -Q^\top A^\top Q \end{bmatrix} \begin{bmatrix} R & \\ & R \end{bmatrix} \\
&= \left[\begin{array}{cc} R_{11}^\top & 0 \\ R_{12}^\top & 0 \\ \hline 0 & R_{11}^\top \\ 0 & R_{12}^\top \end{array}\right] \begin{bmatrix} \widetilde{\Phi}_k & -\widetilde{G}_{11} \\ -\widetilde{H}_{11} & -\widetilde{\Phi}_k^\top \end{bmatrix} \left[\begin{array}{cc|cc} R_{11} & R_{12} & 0 & 0 \\ 0 & 0 & R_{11} & R_{12} \end{array}\right] + O(\epsilon).
\end{aligned}$$

So $\widetilde{\mathcal{H}}$ is the sum of a low-rank matrix and an $O(\epsilon)$ perturbation thus has some $O(\epsilon)$ eigenvalues, rendering it *nearly* non-Hamiltonian and the corresponding pCARE not uniquely solvable. This illustrates the importance of truncation.

4 Numerical Examples

We present four examples to illustrate the feasibility of the projection method with $\mathcal{K}_k(A_\gamma^{-\top},\, C^\top)$. The non-decreasing quantities in the first set of inequalities in (30) are listed in the 3rd–5th columns in the tables, where k is the index in $\mathcal{K}_k$ and the normalized residual is denoted by

$$\text{NRes} \equiv \frac{\|R_k\|}{2\|A^\top V_k Y_k\| + \|Y_k G_{11} Y_k\| + \|H_{11}\|}.$$

Also, ρ_k, $\|r_k\|$ and $\kappa(Y_k)$ are defined as in (5) and (30), "Rank" is the rank of V_k, dt_k the execution time consumed in the kth iteration and $t_k \equiv \sum_{j\leq k} dt_j$. We terminate our computation when NRes < tol_ck. Recall that we may truncate under the control of tol_def when applying the QR decomposition in the Arnoldi process, as discussed in Section 3.1. Importantly, the tables display the numerical results as if the projection method is iterative. In reality, we start with some k then increase it to $\tilde{k}$ until some desired accuracy is achieved. We may extrapolate R_k with respect to k when choosing $\tilde{k}$.

We choose the shift γ in our rational Krylov subspace [13, 14, 15, 16, 20, 21, 43, 46, 48] as in [37]. For our examples, we choose γ arbitrarily as the choice seems unimportant. For a more elaborate strategy, consult [48].

All examples have been attempted using MATLAB Ver. R2013a on an HP Z420 with an Intel Xeon CPU E5-16200 at 3.60 GHz and a 32 GB RAM.

4.1 Example 0 (Randomly Generated)

Example 0 is generated randomly to illustrate the case when r_k diminishes with respect to k. Here we have $n = 10000$, $m = 4$, $l = 2$, $s = 10$, $\alpha = 10^{-14}$, $B = \text{rand}(n, m)$, $C = \text{rand}(l, n)$ and

$$A = \begin{bmatrix} -1 & \alpha & & & & \\ \alpha & \cdots & \alpha & & & \\ & \ddots & -s & \ddots & & \\ & & \alpha & -(s+1)\alpha & \alpha & \\ & & & \ddots & \ddots & \alpha \\ & & & & \alpha & -n\alpha \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

We have tol_def $= 0.0$ as there is no need for any truncation.

Table 1: Example 0. $n = 10000$, $m = 4$, $l = 2$, $\gamma = 5$, tol_ck $= 1.00\text{e}{-14}$, $\alpha = 1.00\text{e}{-14}$

k	NRes	ρ_k	$\|r_k\|$	$\kappa(Y_k)\rho_k$	$\kappa(Y_k)$	Rank	dt_k	t_k
1	1.1e−03	2.1e−01	2.9e−01	1.1e+01	5.4e+01	2	8.73e−02	8.73e−02
2	4.8e−05	1.4e−02	1.9e+00	5.0e+04	3.7e+06	4	8.62e−03	9.59e−02
3	3.6e−06	1.3e−03	2.5e+00	1.1e+06	8.1e+08	6	3.91e−03	9.99e−02
4	2.8e−07	1.3e−04	1.7e+00	1.7e+05	1.3e+09	8	3.61e−03	1.03e−01
5	2.0e−08	1.0e−05	2.2e+00	1.9e+06	2.0e+11	10	1.03e−02	1.14e−01
6	1.5e−14	1.4e−11	2.7e−06	2.4e+02	1.7e+13	12	4.59e−03	1.18e−01
7	1.2e−14	1.3e−11	2.6e−11	5.2e+05	4.1e+16	14	7.54e−03	1.26e−01

Here, $\|r_k\|$ is decreasing, with V_k accurate enough and r_k is practically zero for $k = 7$, and NRes approaches near machine accuracy. Only 1.26 seconds of execution time is required for the projection method with $\mathcal{K}_k(A_\gamma^{-\top}, C^\top)$ for $k = 1, \cdots, 7$. With $\mathcal{K}_k(A^\top, C^\top)$, the results are similar to that in Table 1, as described in Remark 2.1. From the 3rd–5th columns in Table 1, the first set of inequalities in (30) is illustrated. Notice also the increasing condition number $\kappa(Y_k)$ as predicted. The times in the last two columns are satisfyingly small. Similar comments hold for Examples 1–3.

4.2 Example 1 (Cooling of Steel Profile)

The example is quoted originally from [6, Chapter 19] with the updated data, in the form of a generalized CARE (GCARE), from [25]:

$$A^\top XE + E^\top XA - E^\top XGXE + H = 0,$$

for some $E \in \mathbb{R}^{n\times n}$. This is equivalent to the CARE:

$$(E^{-1}A)^\top \widetilde{X} + \widetilde{X}(E^{-1}A) - \widetilde{X}(E^{-1}GE^{-\top})\widetilde{X} + H = 0,$$

where $\widetilde{X} \equiv E^\top XE$. To solve the above CARE, $E^{-1}A$ is not required explicitly. Instead, we have

$$\mathcal{K}_k((E^{-1}A - \gamma I)^{-\top}, C^\top) \quad = \quad \mathcal{K}_k(E^\top (A - \gamma E)^{-\top}, C^\top),$$

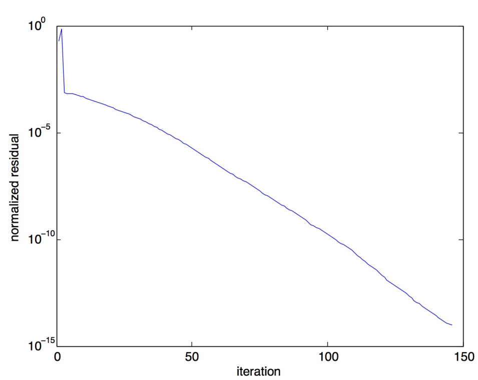
Table 2: Example 1. $n = 79841$, $m = 7$, $l = 6$; $\gamma = 1$, tol_ck = 1.00e−14

k	NRes	ρ_k	$\|r_k\|$	$\kappa(Y_k)\rho_k$	$\kappa(Y_k)$	Rank	dt_k	t_k
1	2.8e−01	4.6e+00	7.3e+00	7.4e+01	1.6e+01	6	1.40e+00	1.40e+00
2	9.3e−02	1.1e+00	1.1e+01	9.4e+01	8.5e+01	12	1.30e+00	2.70e+00
3	3.7e−02	6.2e−01	1.3e+01	1.8e+02	2.9e+02	18	1.44e+00	4.14e+00
4	1.8e−02	3.1e−01	1.4e+01	2.9e+02	9.3e+02	24	1.53e+00	5.68e+00
5	9.4e−03	2.0e−01	1.6e+01	4.1e+02	2.0e+03	30	1.51e+00	7.19e+00
6	5.4e−03	1.2e−01	1.9e+01	8.1e+02	7.0e+03	36	1.47e+00	8.66e+00
7	3.0e−03	7.6e−02	2.0e+01	1.8e+03	2.3e+04	42	1.61e+00	1.03e+01
8	2.1e−03	4.6e−02	2.1e+01	4.1e+03	8.8e+04	48	1.56e+00	1.18e+01
9	1.3e−03	3.3e−02	2.2e+01	9.8e+03	3.0e+05	54	1.70e+00	1.35e+01
10	9.9e−04	2.1e−02	2.2e+01	2.5e+04	1.2e+06	60	1.75e+00	1.53e+01
...								
245	8.6e−15	6.0e−12	8.1e+01	5.5e+06	9.2e+17	1470	7.46e+01	5.65e+03

involving $E^\top$ and $(A-\gamma E)^{-\top}$. (There are other ways to handle GCAREs, outside the scope of this paper.) We have chosen tol_def = 0.0 without truncation. We summarize the numerical results in Table 2 and Figure 1.

For Example 1, $\|r_k\|$ actually increases with respect to k yet NRes approaches near machine accuracy with $k = 1{:}245$ in 5,650 seconds. Notice again the increasing condition number $\kappa(Y_k)$. If we know in advance the required dimension of $\mathcal{K}_k$ for a desired accuracy, much less execution time will be required. For example, for $k = 245$, only 74.6 seconds of execution time is required. Interestingly, the accuracy of the finite element model behind the example has accuracy, at best, around $O(10^{-3})$ and a similar accuracy can be achieved with k =1:5 in 7.19 seconds.

Figure 1: Example 1 (Steel Profile) CARE, $n = 79841$



4.3 Example 2 (Jbilou [26, Example 1])

We quote the example from [26], with $d = 0.5$, $B = \text{rand}(n, 4)$, $C = I_{8\times n}$ and

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

The Krylov subspaces $\mathcal{K}_k(A^\top, C^\top)$ improve slowly with increasing k, changing only the first column in v_{k+1}. The improvement in the Krylov subspace $\mathcal{K}_k(A_\gamma^{-\top}, C^\top)$ is equally slow in our experiments, with $\|r_k\|$ refusing to diminish. We summarize the numerical results in Table 3.

Table 3: Example 2. $n = 1000$, $m = 4$, $l = 8$; $\gamma = 5$, tol_def = 1.00e−10, tol_ck = 1.00e−15

k	NRes	ρ_k	$\|r_k\|$	$\kappa(Y_k)\rho_k$	$\kappa(Y_k)$	Rank	dt_k	t_k
1	6.8e−02	6.1e−01	1.0e+00	1.8e+00	2.9e+00	8	4.99e−02	4.99e−02
2	2.5e−03	3.5e−02	1.5e+00	1.6e+01	4.7e+02	10	7.29e−03	5.72e−02
3	1.4e−04	2.6e−03	1.5e+00	2.9e+02	1.1e+05	12	2.47e−03	5.97e−02
4	9.1e−06	2.1e−04	1.5e+00	5.4e+03	2.6e+07	14	2.53e−03	6.22e−02
5	6.2e−07	1.7e−05	1.5e+00	1.1e+05	6.1e+09	16	2.74e−03	6.49e−02
6	4.5e−08	1.5e−06	1.5e+00	2.1e+06	1.5e+12	18	2.92e−03	6.79e−02
7	3.4e−09	1.3e−07	1.5e+00	4.3e+07	3.5e+14	20	5.23e−03	7.31e−02
8	2.6e−10	1.1e−08	1.5e+00	9.0e+08	8.3e+16	22	7.82e−03	8.09e−02
9	2.0e−11	9.5e−10	1.5e+00	3.1e+10	3.2e+19	24	8.16e−03	8.91e−02
10	1.6e−12	8.3e−11	1.5e+00	9.9e+08	1.2e+19	26	4.39e−03	9.35e−02
11	1.3e−13	7.3e−12	1.5e+00	3.0e+07	4.1e+18	28	5.24e−03	9.87e−02
12	1.1e−14	6.5e−13	1.5e+00	5.2e+06	8.1e+18	30	6.32e−03	1.05e−01
13	8.8e−16	5.8e−14	1.5e+00	3.8e+05	6.7e+18	32	6.02e−03	1.11e−01

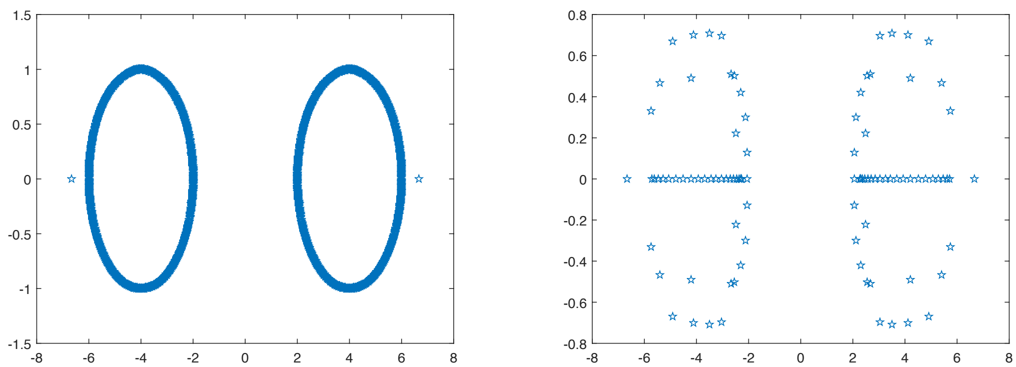
The norms of the Arnoldi residuals $\|r_k\| = 1.5$ stay constant for $k = 2, \cdots, 13$ but the normalized residual NRes achieves the machine accuracy of 4.5e−16 when $k = 13$. However, the Hamiltonian matrices $\mathcal{H}$ and $\widetilde{\mathcal{H}}$ for the original and pCAREs have no eigenvalues near the imaginary axis. The spectral information for Example 2 is summarized in Figure 2. The distances from the imaginary axis for the original CARE and pCARE fall respectively in [1.9937, 6.6567] and [2.0484, 6.6550], so the stability radius has improved slightly after projection.

We have a large stability radius $\tau(A, B) \approx 2 > \|r_k\| \approx 1.5$ ($k \geq 2$), implying an "easy" CARE and the inheritance of stabilizability. The result in (30) and the ill-condition of Y_k are illustrated in the 3rd–6th columns. The ranks of V_k, thus the execution times required, are small. For $\mathcal{K}_k(A^\top, C^\top)$, we achieve the slightly worse result with $k = 22$, NRes = 4.4e−16, $\|r_k\| = 1.5$, Rank = 50 and t_k = 1.92e−01.

Example 2 is originally designed to illustrate the stagnation of r_k. However, from our deeper analysis and numerical experience, $\|r_k\|$ are actually small relative to $\Psi(A - GX)$ for moderate values of k, yielding small R_k's and accurate approximate solutions X_k.

Figure 2: Spectral properties for Example 2

(a) $\Lambda(\mathcal{H})$ (b) $\Lambda(\widetilde{\mathcal{H}})$



4.4 Example 3 (Convective Thermal Flow [23, Example 4.2])

The original model came from [42]. No truncation is applied with tol_def = 0.0. We summarized the results in Table 4 and Figure 3.

Table 4: Example 3. $n = 9669$, $m = 1$, $l = 5$, $\gamma = 1$, tol_ck = 1.00e−14

k	NRes	ρ_k	$\|r_k\|$	$\kappa(Y_k)\rho_k$	$\kappa(Y_k)$	Rank	dt_k	t_k
1	1.9e−01	2.0e+04	2.7e+03	1.0e+05	4.9e+00	5	1.43e−01	1.43e−01
2	7.0e−01	1.3e+01	5.2e+03	5.9e+08	4.6e+07	10	8.49e−02	2.28e−01
3	7.0e−04	6.8e+01	1.1e+04	4.6e+05	6.7e+03	15	7.91e−02	3.07e−01
4	6.9e−04	7.1e+01	1.7e+04	7.9e+05	1.1e+04	20	8.31e−02	3.90e−01
...								
145	1.0e−14	2.3e−09	6.3e+04	3.4e+11	1.5e+20	148	6.70e+00	2.79e+02
146	9.7e−15	2.3e−09	6.0e+04	1.4e+12	6.2e+20	148	6.90e+00	2.86e+02

The Arnoldi residual r_k does not diminishes as the residual R_k. A near machine accuracy for NRes is achieved with $k = 1{:}146$ in 2,860 seconds. Again, for the finite element model, an arguably acceptable accuracy of $O(10^{-3})$ can be achieved with $k = 1{:}3$ in 0.3 second.

5 Conclusions

We have presented additional links between the structure-preserving doubling algorithm and projection methods with rational Krylov subspaces. An analysis of projection methods, when $\|\check{x}_1^\top r_k\|$ or $\|r_k \check{y}_2\|$ small relative to $\tau(A, B)$ or $\tau(A^\top, C^\top)$ respectively, has been presented. The inheritance properties of stabilizability, detectability and other conditions of solvability for projection methods on CAREs have been investigated. This forms a good basis for further research for the inheritance properties of projection methods on CAREs.

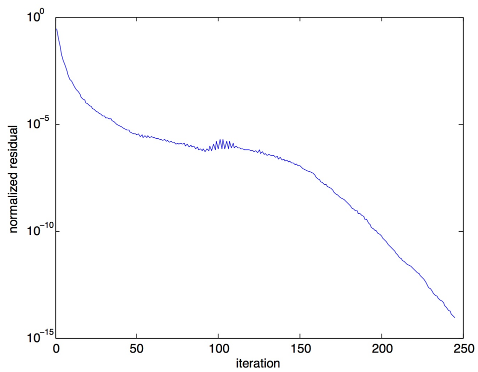


Figure 3: Example 3 (Convective Thermal Flow [23, Example 4.2]) CARE, $n = 9669$

We have presented some numerical results for the projection method using $\mathcal{K}_k(A_\gamma^{-\top}, C^\top)$. A comprehensive comparison of different Krylov subspaces in projection methods for AREs is a worthwhile but large project for the future. We would also love to illustrate numerically the various inheritance properties associated with AREs. However, the estimation of the distances to unstabilzability and undetectability or stability radius is difficult or expensive to realize [19, 24, 28, 30, 31, 33, 34, 44, 45], especially for large-scale problems. Similar studies of inheritance properties for other types of Riccati equations and related linear matrix equations are also possible.

5.1 Applicability of Inheritance Properties

As in most numerical methods, solvability conditions can be difficult or expensive to check. This is also the case for the inheritance properties for projection methods in Section 2.2. This makes some theories somewhat academic, almost impractical. Also, most authors utilize r_k in their theories on projection methods yet it is well known (e.g., from [26] and the examples in Section 4.2) that $\|r_k\|$ may not diminish quickly for moderate values of k. This creates certain amount of difficulties, theoretical and numerical, for projection methods. There seem to be four possibilities:

1. The Arnoldi residual r_k deteriorates in norm with respect to k, thus the conditions $\tau(A, B) > \|\check{x}_1^\top r_k\|$ and $\tau(A^\top, C^\top) > \|r_k \check{y}_2\|$ for the inheritance properties in Theorems 2.2 and 2.3 are satisfied, hopefully, for some *large* enough values of k. This provides a theoretical foundation, alas with conditions unchecked, for the corresponding projection methods.

2. The Arnoldi residual r_k stagnates or persists to be significant in norm with respect to k but still satisfies the conditions of the inheritance properties. Users of the projection methods will not be aware of the solvability of the pCAREs without expensive checks.

3. The Arnoldi residual r_k stagnates but $\|R_k\| = \|Y_k r_k\|$, the norm of the residual for the approximate solution $X_k = P_1 Y_k P_1^\top$, converges to zero quickly. Because of the results in [48] and Sections 2 and 3, the projection method produces an accurate approximate solution X_k. In particular, Theorem 2.5, independent of r_k, may be adequate for most applications.

4. The quantity $\|r_k\|$ stagnates and $\|R_k\|$ remains large until very large values of k, rendering the corresponding projection method infeasible. This may be the fault of an inappropriate Krylov subspace or the solution X is numerically high rank (e.g., when H is high rank). Slow progress in numerical computations should reveal the difficulty and users have to choose better Krylov subspaces or investigate the problem at hand more carefully.

As always, there are aspects on which our understanding is shallow and more research has to be conducted. Still, it is arguably better to have an incomplete theory, in the first step in a long search, than no theory at all. We are far from a thorough understanding of the inheritance properties for projection methods.

Acknowledgements

Part of the work occurred when the first and third authors visited the School of Mathematical Sciences at Fudan University. The first and the second authors were supported by the National Natural Science Foundation, China (Grant 11601484) and the Ministry of Science and Technology, Taiwan (Grant #MOST-105-2115-M-003-003), respectively.

References

[1] L. Amodei and J.-M. Buchot, An invariant subspace method for large-scale algebraic Riccati equation, *Appl. Numer. Maths.*, **60** (2010) 67–1082.

[2] E. Bänsch and P. Benner, Stabilization of incompressible flow problems by Riccati-based feedback, In G. Leugering, S. Engell, A. Griewank, M. Hinze, R. Rannacher, V. Schulz, M. Ulbrich and S. Ulbrich (editors), *Constrained Optimization and Optimal Control for Partial Differential Equations*, International Series of Numerical Mathematics, Vol. **160**, pp. 5–20, Birkhäuser, Basel, 2011.

[3] P. Benner and Z. Bujanović, On the solution of large-scale algebraic Riccati equations by using low-dimensional invariant subspaces, *Lin. Alg. Applic.*, **488** (2016) 430–459.

[4] P. Benner, Z. Bujanović, P. Kürschner and J. Saak, *RADI: A low-rank ADI-type algorithm for large-scale algebraic Riccati equations*, Tech. Rept., 2015; arXiv:1510.00040/v3.

[5] P. Benner, J.-R. Li and T. Penzl, Numerical solution of large Lyapunov equations, Riccati equations and linear-quadratic control problems, *Numer. Lin. Alg. Appl.*, **15** (2008) 755–777.

[6] P. Benner, V. Mehrmann and D. Sorensen, eds., *Dimension Reduction of Large-Scale Systems*, Vol. **45** of Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin/Heidelberg, Germany, 2005.

[7] P. Benner and J. Saak, A Galerkin-Newton-ADI method for solving large-scale algebraic Riccati equations, *DFG Priority Programme 1253 "Optimization with Partial Differential Equations"*, Preprint **SPP1253-090**, January 2010; available at `www.am.uni-erlangen.de/home/spp1253/wiki/images/2/28/Preprint-SPP1253-090.pdf`.

[8] P. BENNER AND J. SAAK, *Numerical Solution of Large and Sparse Continuous Time Algebraic Matrix Riccati and Lyapunov Equations: A State of the Art Survey,* GAMM-Reports, **6** (2013) 32–52.

[9] D.A. BINI, B. IANNAZZO AND B. MEINI, *Numerical Solution of Algebraic Riccati Equations*, SIAM, Philadelphia, 2012.

[10] 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.

[11] E.K.-W. CHU, H.-Y. FAN AND W.-W. LIN, A structure-preserving doubling algorithm for continuous-time algebraic Riccati equations, *Lin. Alg. Appl.*, **396** (2005) 55–80.

[12] E.K.-W. CHU, H.-Y. FAN, W.-W. LIN AND C.-S. WANG, A structure-preserving doubling algorithm for periodic discrete-time algebraic Riccati equations, *Int. J. Control*, **77** (2004) 767–788.

[13] K. DECKERS AND A. BULTHEEL, *Rational Krylov sequences and orthogonal rational functions*, Tech. Rep., Department of Computer Science, KU Leuven, 2007.

[14] 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.

[15] V. DRUSKIN, C. LIEBERMAN AND M. ZASLAVSKY, On adaptive choice of shifts in rational Krylov subspace reduction of evolutionary problems, *SIAM J. Sci. Comput.*, **32** (2010) 2485–2496.

[16] V. DRUSKIN AND V. SIMONCINI, Adaptive rational Krylov subspaces for large-scale dynamical systems, *Syst. Control Letts.*, **60** (2011) 546–560.

[17] L. GIRAUD, J. LANGOU AND M. ROZLOZNIK, The loss of orthogonality in the Gram-Schmidt orthogonalization process, *Computers Maths. Applic.*, **50** (2005) 1069–1075.

[18] G.H. GOLUB AND C.F. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.

[19] M. GU, E. MENGI, M.L. OVERTON, J. XIA AND J. ZHU, Fast methods for estimating the distance to uncontrollability, *SIAM J. Matrix Anal. Applic.*, **28** (2006) 477–502.

[20] S. GÜTTEL, Rational Krylov approximation of matrix functions: numerical methods and optimal pole selection, *GAMM Mitteilungen*, **36** (2013) 8–31.

[21] S. GÜTTEL AND L. KNIZHNERMAN, Automated parameter selection for rational Arnoldi approximation of Markov functions, *Proc. Appl. Math. Mech.*, **11** (2011) 15–18.

[22] C. HE AND G.A. WATSON, An algorithm for computing the distance to instability, *SIAM J. Matrix Anal. Appl.*, **20** (1998) 101–116.

[23] M. HEYOUNI AND K. JBILOU, An extended block Arnoldi algorithm for large-scale solutions of the continuous-time algebraic Riccati equations, *Elect. Trans. Numer. Anal.*, **33** (2009) 53–62.

[24] G. HU AND F.J. DAVISON, Real controllability/stabilizability radius of LTI systems, *IEEE Trans. Automat. Control*, **49** (2004) 254–257.

[25] IMTEK, A Semi-discretized Heat Transfer Problem for Optimal Cooling of Steel Profiles (38881), IMTEK – Simulation, University of Freiburg, 2017; available in `https://portal.uni-freiburg.de/imteksimulation/downloads/benchmark/Steel%20Profiles%20%2838881%29`

[26] K. JBILOU, Block Krylov subspace methods for large continuous-time algebraic Riccati equations, *Numer. Algor.*, **34** (2003) 339–353.

[27] K. JBILOU, An Arnoldi based algorithm for large algebraic Riccati equations, *Appl. Math. Lett.*, **19** (2006) 437–444.

[28] F. KANGAL, K. MEERBERGEN, E. MENGI AND W. MICHIELS, *A subspace method for large scale eigenvalue optimization*, arXiv:1508.0421v2[mathNA], 2015.

[29] C. KENNEY, A.J. LAUB AND M. WETTE, Error bounds for Newton refinement of solutions to algebraic Riccati equations, *Math. Control Signals Syst.*, **3** (1990) 211-224.

[30] C. KENNEY AND A. LAUB, Controllability and stability radii for companion form systems, *Math. Control Signals Syst.*, **1** (1988) 239–256.

[31] D.C. KHANH, H.T. QUYEN AND D.D.X. THANH, On computing stabilizability radii of linear time-invariant continuous systems, *Elect. Trans. Numer. Anal.*, **40** (2013) 407–413.

[32] D. KLEINMAN, On an iterative technique for Riccati equation computations, *IEEE Trans. Automat. Control*, **13** (1968) 114–115.

[33] D. KRESSNER, Deflation in Krylov subspace methods and distance to uncontrollability, *Annali dell'Universitá di Ferrara,* Sezione 7: Scienze matematiche **53** (2007) 309–318; DOI 10.1007/s11565-007-0022-0.

[34] D. KRESSNER, Subspace methods for computing the pseudospectral abscissa and the stability radius, *SIAM J. Matrix Anal. Appl.*, **35** (2014) 292–313.

[35] P. LANCASTER AND L. RODMAN, *Algebraic Riccati Equations*, Clarendon Press, Oxford, 1995.

[36] A.J. LAUB, A Schur method for solving algebraic Riccati equation, *IEEE Trans. Automat. Control*, **AC-24** (1979) 913–921.

[37] T. LI, E.K.-W. CHU, W.-W. LIN AND P.C.-Y. WENG, Solving large-scale continuous-time algebraic Riccati equations by doubling, *J. Comput. Appl. Math.*, **237** (2013) 373–383.

[38] T. LI , C.-Y. WENG, E.K.-W. CHU AND W.-W. LIN, Large-scale Stein and Lyapunov equations, Smith method, and applications, *Numer. Algor.*, **63** (2013) 727–752.

[39] Y. LIN AND V. SIMONCINI, A new subspace iteration method for the algebraic Riccati equation, *Numer. Lin. Alg. Appl.*, **22** (2015) 26–47.

[40] A. MASSOUDI, M.R. OPMEER AND T. REIS, Analysis of an iteration method for the algebraic Riccati equation, *SIAM J. Matrix Anal. Applic.*, **37** (2016) 624–648.

[41] V.L. MEHRMANN, *The Autonomous Linear Quadratic Control Problem*, Lecture Notes in Control and Information Sciences, Vol. **163**, Springer Verlag, Berlin, 1991.

[42] C. MOOSMANN, E.B. RUDNYI, A. GREINER AND J.G. KORVINK, Model order reduction for linear convective thermal flow, in *Proceedings of 10th International Workshops on THERMal INvestigations of ICs and Systems*, THERMINIC2004, Sophia Antipolis, France, Sept. 2004, pp. 317–321.

[43] K.H.A. OLSSON AND A. RUHE, Rational Krylov for eigenvalue computation and model order reduction, *BIT Numer. Math.*, **46** (2006) 99–111.

[44] C.C. PAIGE, Properties of numerical algorithms relating to computing controllability, *IEEE Trans. Automat. Control*, **26** (1981) 130–138.

[45] R. RAJAMANI AND Y.M. CHO, On stability radius and state feedback, *Control Intelligent Syst.*, **32** (2004) 99–103.

[46] A. RUHE, Rational Krylov sequence methods for eigenvalue computation, *Lin. Alg. Appl.*, **58** (1984) 391–405.

[47] J. SAAK, H. MENA AND P. BENNER, *Matrix Equation Sparse Solvers (MESS): A Matlab Toolbox for the Solution of Sparse Large-Scale Matrix Equations*, Chemnitz University of Technology, Germany, 2010.

[48] 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.

[49] V. SIMONCINI, D. B. SZYLD AND M. MONSALVE, On two numerical methods for the solution of large-scale algebraic Riccati equations, *IMA J. Numer. Anal.*, **34** (2014) 904–920.

[50] G.W. STEWART, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, *SIAM Rev.*, **15** (1973) 752–764.

[51] J.-G. SUN, Perturbation theory for algebraic Riccati equations, *SIAM J. Matrix Anal. Applic.*, **19** (1998) 39–65.

[52] C. VAN LOAN, How near is a stable matrix to an unstable matrix? *Contemp. Maths.*, **47** (1985) 465–478.