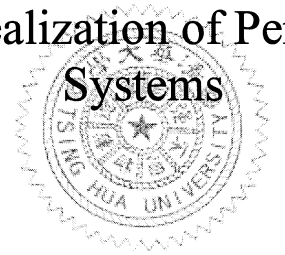
國立清華大學數學系博士班

博士論文

代數黎卡迪方程式之數值研究與週期奇異系統之平衡實現化理論

Numerical Study of Algebraic Riccati Equations and Balanced Realization of Periodic Descriptor Systems



指導教授：林文偉 (Wen-Wei Lin)

研 究 生：范洪源 (Hung -Yuan Fan)

學　　號：887206

中 華 民 國 九 十 三 年　六 月

國 立 清 華 大 學

博碩士論文全文電子檔著作權授權書

（提供授權人裝訂於紙本論文書名頁之次頁用）

本授權書所授權之學位論文，為本人於國立清華大學數學系
_____組， 92 學年度第 2 學期取得博士學位之論文。

論文題目：代數黎卡迪方程式之數值研究與週期奇異系統之平衡實現化理論
指導教授：林文偉

■ 同意

本人茲將本著作，以非專屬、無償授權國立清華大學與台灣聯合大學系統圖書館：基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會與學術研究之目的，國立清華大學及台灣聯合大學系統圖書館得不限地域、時間與次數，以紙本、光碟或數位化等各種方法收錄、重製與利用；於著作權法合理使用範圍內，讀者得進行線上檢索、閱覽、下載或列印。

論文全文上載網路公開之範圍及時間：

本校及台灣聯合大學系統區域網路	■ 立即公開
校外網際網路	■ 立即公開

■ **全文電子檔送交國家圖書館**

授 權 人：范洪源

親筆簽名： 范洪源

中華民國 93 **年** 6 **月** 30 **日**

國 立 清 華 大 學

博碩士紙本論文著作權授權書

（提供授權人裝訂於全文電子檔授權書之次頁用）

本授權書所授權之學位論文，為本人於國立清華大學數學系
_____組， 92 學年度第 2 學期取得博士學位之論文。

論文題目：代數黎卡迪方程式之數值研究與週期奇異系統之平衡實現化理論
指導教授：林文偉

■ 同意

本人茲將本著作，以非專屬、無償授權國立清華大學，基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會與學術研究之目的，國立清華大學圖書館得以紙本收錄、重製與利用；於著作權法合理使用範圍內，讀者得進行閱覽或列印。

本論文為本人向經濟部智慧局申請專利(未申請者本條款請不予理會)的附件之一，申請文號為：_____，請將論文延至____年____月____日再公開。

授 權 人：范洪源

親筆簽名： 范洪源

中華民國 93 年 6 月 30 日

國家圖書館
博碩士論文電子檔案上網授權書

（提供授權人裝訂於紙本論文本校授權書之後）

ID:GH000887206

本授權書所授權之論文為授權人在國立清華大學數學系 92 學年度第 2 學期取得博士學位之論文。

論文題目：代數黎卡迪方程式之數值研究與週期奇異系統之平衡實現化理論
指導教授：林文偉

茲同意將授權人擁有著作權之上列論文全文（含摘要），非專屬、無償授權國家圖書館，不限地域、時間與次數，以微縮、光碟或其他各種數位化方式將上列論文重製，並得將數位化之上列論文及論文電子檔以上載網路方式，提供讀者基於個人非營利性質之線上檢索、閱覽、下載或列印。

☐ 讀者基於非營利性質之線上檢索、閱覽、下載或列印上列論文，應依著作權法相關規定辦理。

授權人：范洪源

親筆簽名： 范洪源

民國 93 年 6 月 30 日

博碩士論文授權書

(國科會科學技術資料中心版本，93.2.6)

本授權書所授權之論文爲本人在__國立清華__大學(學院)____數學系__系所
_____組____92_____學年度第__2__學期取得__博__士學位之論文。

論文名稱：代數黎卡迪方程式之數值研究與週期奇異系統之平衡實現化理論

☑ 同意　　☐不同意

本人具有著作財產權之論文全文資料，授予行政院國家科學委員會科學技術資料中心(或其改制後之機構)、國家圖書館及本人畢業學校圖書館，得不限地域、時間與次數以微縮、光碟或數位化等各種方式重製後散布發行或上載網路。
本論文爲本人向經濟部智慧財產局申請專利（未申請者本條款請不予理會）的附件之一，申請文號爲：_____，註明文號者請將全文資料延後半年後再公開。

☑ 同意　　☐不同意

本人具有著作財產權之論文全文資料，授予教育部指定送繳之圖書館及本人畢業學校圖書館，爲學術研究之目的以各種方法重製，或爲上述目的再授權他人以各種方法重製，不限地域與時間，惟每人以一份爲限。

上述授權內容均無須訂立讓與及授權契約書。依本授權之發行權爲非專屬性發行權利。依本授權所爲之收錄、重製、發行及學術研發利用均爲無償。上述同意與不同意之欄位若未鉤選，本人同意視同授權。

指導教授姓名：林文偉

研究生簽名：范洪源　學號：887206
（親筆正楷）　（務必填寫）

日期：民國 93 年 6 月 17 日

國立清華大學博士學位論文
指導教授推薦書

數學 系 范洪源 君所提之論文

代數黎卡迪方程式之數值研究與
週期奇異系統之平衡實現化理論

Numerical Study of Algebraic Riccati Equations
and Balanced Realization of Periodic
Descriptor Systems

經由本人指導撰述，同意提付審查。

指導教授 林文偉 (簽章)

中華民國 93 年 6 月 17 日

國立清華大學博士學位論文

考試委員審定書

數學 系 范洪源 君所提之論文

代數黎卡迪方程式之數值研究與
週期奇異系統之平衡實現化理論

Numerical Study of Algebraic Riccati Equations and Balanced Realization of Periodic Descriptor Systems

經本委員會審查，符合博士資格標準。

學位考試委員會

主持人 林松山 (簽章)

委 員 王振男 范洪源 葉芳柏 石至文 許文偉

中華民國 93 年 6 月 17 日

中文摘要

本篇論文主要包括兩部分。第一部分闡述如何運用保結構演算法來求解各種類型的黎卡迪方程式，第二部分主要著眼於週期奇異系統的平衡實現化理論。

在第一部分中，我們分別探討求解週期離散型、連續型以及廣義離散型代數黎卡迪方程式之保結構算法。上述各類算法均以求解離散型代數黎卡迪方程式之保結構演算法為基石，加以推廣而得之。並且我們可在比可穩定化與可偵測化更弱的假設條件下，更進一步證明此一保結構算法的二次收斂性。通過大量 Matlab 測試集的檢驗，可知此一保結構算法無論在精確度上與執行效率上均優於其他算法。

在第二部分中，我們先針對週期奇異系統的完全可達性與完成可觀性，給出一系列的充分必要條件。由這些數學等價條件中，我們可定義出週期可達性與可觀性之葛雷米矩陣，並且可進一步證明出這些對稱半正定的葛雷米矩陣滿足某些廣義週期離散型李雅普諾夫方程式。此外，我們還提出一套數值上穩定且可行的算法來求解這些李雅普諾夫方程式。最後，我們還提出週期奇異系統的平衡實現化問題並且提供解決方案。

Abstract

This dissertation is consisted of two parts. The first part treats of applications of the structure-preserving doubling algorithm (SDA) to solve various algebraic Riccati equations, while the second part concerns with the problem of balanced realization for discrete-time periodic descriptor systems.

In the first part, we investigate structure-preserving algorithms for computing the symmetric positive semi-definite solutions to the periodic discrete-time algebraic Riccati equations (P-DAREs), continuous-time algebraic Riccati equations (CAREs) and generalized discrete-time algebraic Riccati equations (G-DAREs), respectively. All are based on the SDA algorithm for solving the discrete-time algebraic Riccati equations (DAREs). In Section 2 of Chapter 1, we develop the SDA algorithm from a new point of view and show its quadratic convergence under assumptions which are weaker than stabilizability and detectability. With several numerical results, the algorithm is shown to be efficient, out-performing other algorithms on a large set of benchmark problems.

In the second part, necessary and sufficient conditions are derived for complete reachability and observability of periodic time-varying descriptor systems. Applying these conditions, the symmetric positive semi-definite reachability/observability Gramians are defined and can be shown to satisfy some projected generalized discrete-time periodic Lyapunov equations. We propose a numerical method for solving these projected Lyapunov equations, and an illustrative numerical example is given. As an application of our results, the balanced realization of periodic descriptor systems is discussed.

誌謝辭

本篇論文得以順利完成，首先必須感謝我的指導老師林文偉教授，在碩士、博士期間日夜晨昏地對我的殷殷教誨與指導規劃，讓我奠定扎實的研究基礎，完成了這一系列的研究主題。其次，要感謝澳洲Monash大學朱景華教授與王辰樹學長對論文提供增修潤飾的幫助，讓我的論文寫作能力能與日俱增。另外，也要感謝比利時的Paul Van Dooren教授，提攜後進，熱切討論，讓我的研究方向更開闊擴展。最後，我要感謝在學期間的學長與同儕學弟間的殷切討論，使得研究路上，順遂平穩。

此外，感謝我的父母雙親給我的鼓勵與支持，讓我在挫折中感覺到支柱。更感謝我的愛妻，陪我走過這求學的漫漫長路，給予我精神上的力量。謹此，對諸位獻上我最衷心的感激。

Contents

Chapter 1

Structure-Preserving Algorithms for P-DAREs

1 Introduction

In this chapter we investigate structure-preserving algorithms for computing the symmetric positive semi-definite (s.p.s.d.) solutions $\{X_j\}_{j=1}^p$ to the periodic discrete-time algebraic Riccati equations (P-DAREs) of period $p \geq 1$:

$$X_{j-1} = A_j^T X_j A_j - (B_j^T X_j A_j + S_j^T C_j)^T (R_j + B_j^T X_j B_j)^{-1} (B_j^T X_j A_j + S_j^T C_j) + C_j^T Q_j C_j. \tag{1.1}$$

Here, for all j, $A_j = A_{j+p} \in \mathcal{R}^{n_j \times n_{j-1}}$ with $n_j = n_{j+p}$, $X_j = X_{j+p} \in \mathcal{R}^{n_j \times n_j}$, $R_j = R_{j+p} \in \mathcal{R}^{m_j \times m_j}$ and $Q_j = Q_{j+p} \in \mathcal{R}^{r_j \times r_j}$ are symmetric positive definite (or s.p.d.; i.e. $R_j, Q_j > 0$), $B_j = B_{j+p} \in \mathcal{R}^{n_j \times m_j}$, $S_j = S_{j+p} \in \mathcal{R}^{r_j \times m_j}$, and $C_j = C_{j+p} \in \mathcal{R}^{r_j \times n_{j-1}}$, with B_j, C_j^T being of full column rank. Furthermore, the matrix $Q_j - S_j R_j^{-1} S_j^T$ is supposed to be symmetric positive definite. Throughout this chapter, the indices j for all periodic matrices are chosen in $\{1, \cdots, p\}$ modulo p.

Equations in (1.1) arise frequently in the periodic discrete-time linear optimal control problem for the periodic systems

$$\begin{cases} x_{j+1} = A_j x_j + B_j u_j, \quad x_j \in \mathcal{R}^{n_{j-1}}, \\ y_j = C_j x_j \end{cases}$$

with the controls $\{u_j\}$ chosen through optimizing the cost function:

$$\min_{u_j} \mathcal{J} = \frac{1}{2} \sum_{j=1}^{\infty} \left(y_j^T Q_j y_j + u_j^T R_j u_j + y_j^T S_j u_j + u_j^T S_j^T y_j \right).$$

The periodic optimal feedback controls u_j^* are given by [32]

$$u_j^* = -(R_j + B_j^T X_j B_j)^{-1} (B_j^T X_j A_j + S_j^T C_j) x_j \quad (j = 1, \cdots, p) \tag{1.2}$$

where $\{X_j\}_{j=1}^p$ are s.p.s.d. solutions to (1.1).

Definition 1.1. [32]. *The periodic matrix pairs* $\{(A_j, B_j)\}_{j=1}^p$ *are said to be p-stabilizable (P-S) if the pairs* $(\mathcal{A}_j, \mathcal{B}_j)$ *are stabilizable (S), for* $j = 1, \cdots, p$, *where* $\mathcal{A}_j \equiv A_{\pi_j(p)} \cdots A_{\pi_j(1)}$ *and*

$$\mathcal{B}_j \equiv \left[A_{\pi_j(p)} \cdots A_{\pi_j(2)} B_{\pi_j(1)} | A_{\pi_j(p)} \cdots A_{\pi_j(3)} B_{\pi_j(2)} | \cdots | A_{\pi_j(p)} B_{\pi_j(p-1)} | B_{\pi_j(p)}\right]$$

with the permutation π_j *defined by*

$$\pi_j(k) = \begin{cases} k - j + 1 + p\ , & \text{for } k = 1, \cdots, j-1\ , \\ k - j + 1\ , & \text{for } k = j, \cdots, p\ . \end{cases}$$

Definition 1.2. [32]. *The periodic matrix pairs* $\{(A_j, C_j)\}_{j=1}^p$ *are said to be p-detectable (P-D) if the pairs* $(\mathcal{A}_j, \mathcal{C}_j)$ *are detectable (D), for* $j = 1, \cdots, p$, *where* $\mathcal{A}_j$ *and* π_j *are defined as in Definition 1.1, and*

$$\mathcal{C}_j \equiv \left[C_{\pi_j(1)}^T | A_{\pi_j(1)}^T C_{\pi_j(2)}^T | A_{\pi_j(1)}^T A_{\pi_j(2)}^T C_{\pi_j(3)}^T | \cdots | A_{\pi_j(1)}^T \cdots A_{\pi_j(p-1)}^T C_{\pi_j(p)}^T\right]^T\ .$$

Note that the pair (A, B) is stabilizable (S) if $w^T B = 0$ and $w^T A = \lambda w^T$ for some constant λ implies $|\lambda| < 1$ or $w = 0$, and the pair (A, C) is detectable (D) if (A^T, C^T) is stabilizable. Under assumptions of (P-S) and (P-D), P-DAREs have been proved to possess unique s.p.s.d. solutions [31, 32].

Via elementary matrix calculation, one can show that the P-DAREs (1.1) are equivalent to the following form

$$\begin{aligned} X_{j-1} = {} & (A_j - B_j R_j^{-1} S_j^T C_j)^T X_j (A_j - B_j R_j^{-1} S_j^T C_j) \\ & - (A_j - B_j R_j^{-1} S_j^T C_j)^T X_j B_j (R_j + B_j^T X_j B_j)^{-1} B_j^T X_j (A_j - B_j R_j^{-1} S_j^T C_j) \\ & + C_j^T (Q_j - S_j R_j^{-1} S_j^T) C_j, \end{aligned} \tag{1.3}$$

for $j = 1, \ldots, p$. It is easily seen that the periodic matrix pairs $\{(A_j, B_j)\}_{j=1}^p$ are p-stabilizable if and only if $\{(A_j - B_j R_j^{-1} S_j^T C_j, B_j)\}_{j=1}^p$ are p-stabilizable. Similarly, the periodic matrix pairs $\{(A_j, C_j)\}_{j=1}^p$ are p-detectable if and only if $\{(A_j - B_j R_j^{-1} S_j^T C_j, \overline{C}_j)\}_{j=1}^p$ are p-detectable, where $\overline{C}_j^T \overline{C}_j$ is a full rank decomposition (FRD) of $C_j^T (Q_j - S_j R_j^{-1} S_j^T) C_j$ with $\overline{C}_j \in \mathcal{R}^{r_j \times n_{j-1}}$. Consequently, with the following FRDs

$$G_j := B_j R_j^{-1} B_j^T \geq 0, \quad H_j := \overline{C}_j^T \overline{C}_j \geq 0, \tag{1.4}$$

there is no loss of generality to consider, instead of (1.3), the following P-DAREs

$$X_{j-1} = A_j^T X_j A_j - A_j^T X_j B_j (R_j + B_j^T X_j B_j)^{-1} B_j^T X_j A_j + H_j$$

or

$$X_{j-1} = A_j^T X_j (I_{n_j} + G_j X_j)^{-1} A_j + H_j. \tag{1.5}$$

Note that (1.5) is obtained using Sherman-Morrison-Woodbury formula (SMWF; see, e.g., [59, p. 50]) when $(I_{n_j} + G_j X_j)^{-1}$ exists.

We now consider the periodic matrix pairs $\{(M_j, L_j)\}_{j=1}^p$ associated with the P-DAREs in (1.5) with

$$M_j = \begin{bmatrix} A_j & 0 \\ -H_j & I_{n_{j-1}} \end{bmatrix} \in \mathcal{R}^{(n_{j-1}+n_j)\times 2n_{j-1}}, \quad L_j = \begin{bmatrix} I_{n_j} & G_j \\ 0 & A_j^T \end{bmatrix} \in \mathcal{R}^{(n_{j-1}+n_j)\times 2n_j}, \tag{1.6}$$

where I_{n_j} denotes the identity matrix of compatible order. It is easily seen that the matrix pair $(M_j,\ L_j)$ is symplectic, that is,

$$M_j J_j M_j^T = L_j J_{j+1} L_j^T, \quad J_j \equiv \begin{bmatrix} 0 & I_{n_{j-1}} \\ -I_{n_{j-1}} & 0 \end{bmatrix}. \tag{1.7}$$

The matrix pair in the form of (1.6), with $H_j,\ G_j$ being s.p.s.d., is said to be a standard symplectic form (SSF). Being in SSF is a central concept in this chapter and is stronger than being a symplectic form as defined in (1.7). Being in SSF is the structure we try to preserve in the numerical algorithm.

From (1.6), the P-DAREs in (1.5) can be written, for all j, as

$$M_j \begin{bmatrix} I_{n_{j-1}} \\ X_{j-1} \end{bmatrix} = L_j \begin{bmatrix} I_{n_j} \\ X_j \end{bmatrix} \Phi_j \tag{1.8}$$

for some appropriate $\Phi_j \in \mathcal{R}^{n_j \times n_j}$. In the case when all A_j have equal size and are nonsingular, the s.p.s.d. solutions X_j to (1.5) can be easily obtained through the invariant subspaces, associated with the eigenvalues inside the unit disk, of the periodic matrices [62]

$$\Pi_j = L_{j+p-1}^{-1} M_{j+p-1} L_{j+p-2}^{-1} M_{j+p-2} \cdots L_j^{-1} M_j. \tag{1.9}$$

Under the (P-S) and (P-D) assumptions, each Π_j has exactly n such stable eigenvalues. If the columns of $\begin{bmatrix} Z_{1j} \\ Z_{2j} \end{bmatrix}$ spans the stable invariant subspace of Π_j, then Z_{1j} is nonsingular and $X_j = Z_{2j} Z_{1j}^{-1}$ for $j = 1, \cdots, p$. Note that these relations still hold in a generalized sense if some of the A_j are singular or not squared [19, 20, 32, 62, 120]. The theory and algorithms for this general case will be considered in Section 3.

Periodic linear systems arise naturally from continuous linear systems, when multi-rate sampling is performed [55]. These systems have many interesting and practical applications, with notable examples such as the helicopter ground resonance damping problem and the satellite altitude control problems [25, 30, 126]. Large state-space dimensions or large periods appear in different circumstances. The analysis and design of such systems have received much attention in recent years [30, 32, 105, 107, 122, 126]. A numerically backward-stable periodic QZ algorithm for the P-DAREs, which relies on an extension of the generalized Schur method, has been proposed in [33, 62]. Reliable parallel algorithms for solving the P-DAREs based on the swap and collapse technique have been developed in [18, 19, 20, 23, 24, 74].

For the case of $p = 1$, the P-DAREs (1.1) become a single DARE. A well-known backward stable approach, utilizing the QZ algorithm for computing the unique s.p.s.d. solution to a DARE, has been proposed in [88, 96, 119]. Algorithms using symplectic orthogonal transformations for solving DAREs have been proposed in [2, 90]. The doubling algorithms with second-order convergence have been developed in [3, 69]. Matrix sign function-type methods, which solves DAREs implicitly by transforming the symplectic pair into a Hamiltonian matrix, have been developed in [84, 85]. More recently, a matrix disk function method has been developed in [18, 20] based on an inverse-free iteration [7, 86] for computing the unique s.p.s.d. solution of DAREs while preserving the symplectic structure (1.7) in each iterative step.

The QZ-type algorithms [33, 62, 88, 96, 119] (Periodic QZ or QZ) are numerically backward stable, but do not take into account the symplectic structure of (M_j, L_j). Non-structure-preserving iterative processes loosen the symplectic structure, thus may cause

the algorithms to fail or to lose accuracy in adverse circumstances. This will be more serious for ill-conditioned problems, when errors corrupt the stabilizing invariant subspaces and the solution process based on it. The inversion of some potentially ill-conditioned matrices cannot be avoided in the matrix sign function-type methods [84, 85], leading to possible loss of accuracy. The symplectic structure in the algorithms in [2, 90] is preserved only for systems with single input or output. For the general case, the symplectic structure is only retained in exact arithmetic. Similarly, in the matrix disk function/inverse-free methods [7, 18, 19, 20, 23, 24], the symplectic structure can only be preserved in exact arithmetic. The aforementioned problems in non-structure-preserving algorithms will still occur, probably to a lesser extent.

In this chapter, we first revisit the doubling algorithm [3, 69] for solving DAREs while keeping the associated symplectic matrix pairs in SSF in each iterative step. This algorithm attracted much attention but somehow went out of favor in the last decade. We develop the doubling algorithm from a new point of view, which is referred as to the structure-preserving doubling algorithm (SDA), and show the quadratic convergence of the SDA under assumptions which are weaker than (S) and (D). More details can be found in Section 2. Second, we develop a structure-preserving swap and collapse algorithm (SSCA) to reduce the P-DAREs to a single DARE while keeping the associated symplectic matrix pairs in SSF. The P-DAREs can then be solved via the single DARE by SDA.

This chapter is organized as follows. In Section 2, we revisit the doubling algorithm for solving a single DARE, based on the disk function approach [18]. Convergence and error analysis are also presented. The relationship between the disk function method and the doubling algorithm will be discussed. Section 3 contains a structure-preserving algorithm which swaps and collapses the associated symplectic matrix pairs as in (1.6) to a single matrix pair in SSF. In Section 4, we report some numerical results for DAREs selected from [21, 27, 60, 82, 97, 99], comparing the SDA algorithm with the disk function/inverse-free methods [7, 18, 20] and the method associated with `dare` in the MATLAB control toolbox [88]. Section 5 reports the numerical performance of the SSCA+SDA for P-DAREs sampled from [62, 100, 126]. Concluding remarks are given in Section 6.

2 Structure-Preserving Doubling Algorithm for DAREs

Let

$$M = \begin{bmatrix} A & 0 \\ -H & I \end{bmatrix}, \quad L = \begin{bmatrix} I & G \\ 0 & A^T \end{bmatrix} \tag{2.1}$$

where $A \in \mathcal{R}^{n\times n}$, $R \in \mathcal{R}^{m\times m}$ is s.p.d., $B \in \mathcal{R}^{n\times m}$ and $C^T \in \mathcal{R}^{n\times r}$ are of full column rank, $G = BR^{-1}B^T \geq 0$ and $H = C^TC \geq 0$. The pairs (A, B) and (A, C) are assumed to be stabilizable (S) and detectable (D), respectively. Then the DARE

$$X = A^TXA - A^TXB(R + B^TXB)^{-1}B^TXA + H$$

or

$$X = A^TX(I + GX)^{-1}A + H \tag{2.2}$$

has a unique s.p.s.d. solution [96]. In this Section, we apply a swap and collapse procedure to derive the structure-preserving doubling algorithm (SDA) for solving the DARE (2.2), and prove the quadratic convergence of the SDA. Note that the quadratic convergence of the doubling algorithm is proven in [69] for (A, G, H) which is stabilizable and detectable. Below, in Theorem 2.2 we shall prove the quadratic convergence of the SDA under weaker conditions.

Given M and L as in (2.1), we construct

$$T^{(1)} = \left[\begin{array}{cc|cc} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ \hline A & 0 & I & 0 \\ -H & 0 & 0 & I \end{array}\right], \; T^{(2)} = \left[\begin{array}{cc|cc} I & 0 & 0 & 0 \\ 0 & I & 0 & A^T(I+HG)^{-1} \\ \hline 0 & 0 & I & AG(I+HG)^{-1} \\ 0 & 0 & 0 & I \end{array}\right], \; T^{(3)} = \left[\begin{array}{cc|cc} I & 0 & 0 & 0 \\ 0 & 0 & 0 & I \\ \hline 0 & 0 & I & 0 \\ 0 & I & 0 & 0 \end{array}\right]. \tag{2.3}$$

and

$$T \equiv T^{(3)}T^{(2)}T^{(1)} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} I & 0 & 0 & 0 \\ -H & 0 & 0 & I \\ \hline A(I+GH)^{-1} & 0 & I & AG(I+HG)^{-1} \\ -A^T(I+HG)^{-1}H & I & 0 & A^T(I+HG)^{-1} \end{array}\right]. \tag{2.4}$$

We then have

$$T \begin{bmatrix} L \\ -M \end{bmatrix} = \left[\begin{array}{cc} I & G \\ 0 & -(I+HG) \\ \hline 0 & 0 \\ 0 & 0 \end{array}\right]. \tag{2.5}$$

The transformation T represents row operations on $\begin{bmatrix} L \\ -M \end{bmatrix}$ and is obtained as follows:

(1) Use the identity matrix in the (1,1)-block in L to annihilate submatrices beneath it.

(2) Then use the resulting (4,2)-block $[-(I+HG)]$ to eliminate the (2,2)- and (3,2)-blocks.

(3) Permute the row-blocks to the block-upper triangular form on the right-hand-side in (2.5).

Ignoring how T is constructed, the above factorization (2.5) can easily be checked by direct multiplication with the help of the SMWF.

From (2.4), we define

$$\widetilde{M} \equiv T_{21} = \begin{bmatrix} A(I+GH)^{-1} & 0 \\ -A^T(I+HG)^{-1}H & I \end{bmatrix}, \quad \widetilde{L} \equiv T_{22} = \begin{bmatrix} I & AG(I+HG)^{-1} \\ 0 & A^T(I+HG)^{-1} \end{bmatrix}, \tag{2.6}$$

and consequently deduce that

$$\widetilde{M}L = \widetilde{L}M. \tag{2.7}$$

We then compute $\widetilde{L}L$ and $\widetilde{M}M$ and apply the SMWF to produce

$$\widehat{L} \equiv \begin{bmatrix} I & \widehat{G} \\ 0 & \widehat{A}^T \end{bmatrix} = \widetilde{L}L \quad \text{and} \quad \widehat{M} \equiv \begin{bmatrix} \widehat{A} & 0 \\ -\widehat{H} & I \end{bmatrix} = \widetilde{M}M, \tag{2.8}$$

where

$$\widehat{A} = A(I + GH)^{-1}A, \tag{2.9}$$

$$\widehat{G} = G + AG(I + HG)^{-1}A^T, \tag{2.10}$$

$$\widehat{H} = H + A^T(I + HG)^{-1}HA \tag{2.11}$$

with \hat{·} denoting the result of one iterative step. Then by (2.8), $(\widehat{M}, \widehat{L})$ is again in SSF and satisfies

$$\widehat{M}^{-1}\widehat{L} = (M^{-1}L)^2 \tag{2.12}$$

provided that M and $\widehat{M}$ are invertible. Otherwise, please refer to the detailed proof in Lemma 2.1 below.

Equations (2.9)–(2.11) have exactly the same form as the doubling algorithm [3, (4)-(5)] (see also the references therein, as well as [69, 74]). However, the original doubling algorithm was derived as an acceleration scheme for the fixed-point iteration from (2.2):

$$X_{k+1} = A^T X_k (I + GX_k)^{-1} A + H.$$

Instead of producing the sequence $\{X_k\}$, the doubling algorithm produces $\{X_{2^k}\}$. Furthermore, the convergence of the doubling algorithm was proven when A is nonsingular [3], and for (A, G, H) which is stabilizable detectable [69]. Our convergence results in Theorem 2.2 are stronger under weaker conditions (which are implied by (S) and (D)). The preservation of stabilizability and detectability is shown in Lemma 2.3. The interesting relationship between the SDA and the swap and collapse procedure in Section 3 is also new. Problems arising from R being ill-conditioned are tackled in [44].

We now describe the SDA for solving the DARE.

SDA Algorithm

Input : A, G, H; τ (a small tolerance);

Output : s.p.s.d. solution X for DARE.

Initialize $j \leftarrow 0$, $A_0 \leftarrow A$, $G_0 \leftarrow G$, $H_0 \leftarrow H$;

Repeat $W \leftarrow I + G_j H_j$,

Solve for $V_1,\ V_2$ **from** $WV_1 = A_j,\ V_2 W^T = G_j$;

$A_{j+1} \leftarrow A_j V_1,\ G_{j+1} \leftarrow G_j + A_j V_2 A_j^T,\ H_{j+1} \leftarrow H_j + V_1^T H_j A_j$;

Stop when $\|H_{j+1} - H_j\|_F \le \tau \|H_{j+1}\|_F$;

Set $X \leftarrow H_{j+1}$.

End of SDA Algorithm

Convergence of SDA

Let $M = \begin{bmatrix} A & 0 \\ -H & I \end{bmatrix}$, $L = \begin{bmatrix} I & G \\ 0 & A^T \end{bmatrix}$, where $G = G^T$, $H = H^T$. Suppose $M - \lambda L$ has no eigenvalues on the unit circle and there exist nonsingular Q, Z such that

$$QMZ = \begin{bmatrix} J_s & 0 \\ 0 & I \end{bmatrix}, \quad QLZ = \begin{bmatrix} I & 0 \\ 0 & J_s \end{bmatrix} \tag{2.13}$$

where the spectrum $\lambda(J_s) \in O_s \equiv \{\lambda :\ |\lambda| < 1\}$. For the convergence analysis, we first prove the following Lemma.

Lemma 2.1. *Let T be any $4n \times 4n$ nonsingular matrix such that*

$$T\left\{\lambda \begin{bmatrix} L & 0 \\ -M & L \end{bmatrix} - \begin{bmatrix} 0 & M \\ 0 & 0 \end{bmatrix}\right\} = \lambda \begin{bmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{bmatrix} - \begin{bmatrix} 0 & M_{12} \\ 0 & M_{22} \end{bmatrix}. \tag{2.14}$$

Then

(i) the pencil $M_{22} - \lambda L_{22}$ is uniquely determined up to a left transformation.

(ii) The pencil $\widehat{M} - \lambda\widehat{L}$ is equivalent to the pencil

$$\Gamma \left\{ \begin{bmatrix} J_s^2 & 0 \\ 0 & I \end{bmatrix} - \lambda \begin{bmatrix} I & 0 \\ 0 & J_s^2 \end{bmatrix} \right\} Z^{-1}$$

where $\widehat{L} \equiv \widetilde{L}L$, $\widehat{M} \equiv \widetilde{M}M$ are given by (2.8), for some nonsingular matrix Γ.

Proof. (i) Partition $T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$. Since $M - \lambda L$ is regular, so is $\lambda \begin{bmatrix} L & 0 \\ -M & L \end{bmatrix} - \begin{bmatrix} 0 & M \\ 0 & 0 \end{bmatrix}$, implying that $\begin{bmatrix} L \\ -M \end{bmatrix}$ is of full column rank. An inspection of (2.14) indicates that the rows of $[T_{21},\ T_{22}]$ form a basis of the null space of $\begin{bmatrix} L \\ -M \end{bmatrix}$. Therefore, the pencil $M_{22} - \lambda L_{22} = T_{21}M - \lambda T_{22}L$ is uniquely determined up to a left transformation.

(ii) From (2.6)–(2.8), we have

$$\begin{bmatrix} T_{11} & T_{12} \\ \widetilde{M} & \widetilde{L} \end{bmatrix} \begin{bmatrix} L & 0 \\ -M & L \end{bmatrix} = \begin{bmatrix} T_{11}L - T_{12}M & T_{12}L \\ 0 & \widehat{L} \end{bmatrix}$$

$$\begin{bmatrix} T_{11} & T_{12} \\ \widetilde{M} & \widetilde{L} \end{bmatrix} \begin{bmatrix} 0 & M \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & T_{11}M \\ 0 & \widehat{M} \end{bmatrix}$$

where $\widetilde{M}$, $\widetilde{L}$ are given in (2.6). From the definition of T in (2.4), we have

$$T_{11} = \begin{bmatrix} I & 0 \\ -H & 0 \end{bmatrix}, \quad T_{12} = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}.$$

Routine manipulations show that

$$T_{11}L - T_{12}M = \begin{bmatrix} I & G \\ 0 & -(I + HG) \end{bmatrix}, \quad T_{11}M = \begin{bmatrix} A & 0 \\ -HA & 0 \end{bmatrix}.$$

Recall that $J \equiv \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$. From (2.4), (2.5) and (2.13), we can choose $\Theta = \begin{bmatrix} J_s & 0 \\ 0 & 0 \end{bmatrix}$

so that

$$T^{(3)}\begin{bmatrix} I_{2n} & J^T\Theta J \\ 0 & I_{2n} \end{bmatrix}\begin{bmatrix} I_{2n} & 0 \\ \Theta & I_{2n} \end{bmatrix}\begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix}\begin{bmatrix} L & 0 \\ -M & L \end{bmatrix}\begin{bmatrix} Z & 0 \\ 0 & Z \end{bmatrix} = \left[\begin{array}{cc|cc} I & 0 & 0 & 0 \\ 0 & -I & 0 & J_s \\ \hline 0 & 0 & I & 0 \\ 0 & 0 & 0 & J_s^2 \end{array}\right],$$

$$T^{(3)}\begin{bmatrix} I_{2n} & J^T\Theta J \\ 0 & I_{2n} \end{bmatrix}\begin{bmatrix} I_{2n} & 0 \\ \Theta & I_{2n} \end{bmatrix}\begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix}\begin{bmatrix} 0 & M \\ 0 & 0 \end{bmatrix}\begin{bmatrix} Z & 0 \\ 0 & Z \end{bmatrix} = \left[\begin{array}{cc|cc} 0 & 0 & J_s & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & J_s^2 & 0 \\ 0 & 0 & 0 & I \end{array}\right].$$

By (i) we have

$$\widehat{M} - \lambda\widehat{L} = \Gamma\left\{\begin{bmatrix} J_s^2 & 0 \\ 0 & I \end{bmatrix} - \lambda\begin{bmatrix} I & 0 \\ 0 & J_s^2 \end{bmatrix}\right\} Z^{-1}$$

for some nonsingular matrix Γ. □

We now prove the following convergence theorems.

Theorem 2.2. *Let* $M = \begin{bmatrix} A & 0 \\ -H & I \end{bmatrix}$, $L = \begin{bmatrix} I & G \\ 0 & A^T \end{bmatrix}$, *where* $G = G^T$, $H = H^T$. *Suppose* $M - \lambda L$ *has no eigenvalues on the unit circle and there exist nonsingular* Q, Z *such that (2.13) holds. Denote* $Z = \begin{bmatrix} Z_1 & Z_3 \\ Z_2 & Z_4 \end{bmatrix}$, $Z_i \in \mathcal{R}^{n\times n}$ *for* $i = 1, 2, 3, 4$. *If* Z_1 *and* Z_4 *are invertible, then the sequences* $\{A_k,\ H_k,\ G_k\}$ *computed by the SDA algorithm satisfy*

(i) $\|A_k\| = O(\|J_s^{2^k}\|) \to 0$ *as* $k \to \infty$,

(ii) $H_k \to X$, *where* X *solves the DARE (2.2):*

$$X = A^T X(I + GX)^{-1}A + H,$$

(iii) $G_k \to Y$, *where* Y *solves the dual DARE*

$$Y = AY(I + HY)^{-1}A^T + G. \tag{2.15}$$

Moreover, the convergence rate in (i)–(iii) above is $O\left(|\lambda_n|^{2^k}\right)$, *where* $|\lambda_1| \le \cdots \le |\lambda_n| < 1 < |\lambda_n|^{-1} \le \cdots \le |\lambda_1|^{-1}$ *with* λ_i, λ_i^{-1} *being the eigenvalues of* $M - \lambda L$ *(including* 0 *and* ∞*).*

Proof. Let $M_0 = M = \begin{bmatrix} A_0 & 0 \\ -H_0 & I \end{bmatrix}$, $L_0 = L = \begin{bmatrix} I & G_0 \\ 0 & A_0^T \end{bmatrix}$. Then

$$M_k \equiv \begin{bmatrix} A_k & 0 \\ -H_k & I \end{bmatrix} = \begin{bmatrix} A_{k-1}(I + G_{k-1}H_{k-1})^{-1}A_{k-1} & 0 \\ -\left[H_{k-1} + A_{k-1}^T(I + H_{k-1}G_{k-1})^{-1}H_{k-1}A_{k-1}\right] & I \end{bmatrix} \tag{2.16}$$

and

$$L_k \equiv \begin{bmatrix} I & G_k \\ 0 & A_k^T \end{bmatrix} = \begin{bmatrix} I & G_{k-1} + A_{k-1}G_{k-1}(I + H_{k-1}G_{k-1})^{-1}A_{k-1}^T \\ 0 & A_{k-1}^T(I + H_{k-1}G_{k-1})^{-1}A_{k-1}^T \end{bmatrix}. \tag{2.17}$$

From Lemma 2.1(ii) and the SDA, we have

$$M_k - \lambda L_k = \Gamma_k \left\{ \begin{bmatrix} J_s^{2^k} & 0 \\ 0 & I \end{bmatrix} - \lambda \begin{bmatrix} I & 0 \\ 0 & J_s^{2^k} \end{bmatrix} \right\} Z^{-1} \tag{2.18}$$

where $\Gamma_k = \begin{bmatrix} \Gamma_{1k} & \Gamma_{3k} \\ \Gamma_{2k} & \Gamma_{4k} \end{bmatrix}$, $k = 1, 2, \ldots$, are suitable nonsingular matrices. Let

$$X = Z_2 Z_1^{-1}, \quad Y = -Z_3 Z_4^{-1}. \tag{2.19}$$

From (2.13), it follows that the spans $\Re\left\{\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}\right\}$ and $\Re\left\{\begin{bmatrix} Z_3 \\ Z_4 \end{bmatrix}\right\}$ respectively form the stable invariant subspaces of $M - \lambda L$ and $(J^T L J) - \lambda(J^T M J)$. By the result of [96], it is clear that the symmetric X and Y solve the DAREs (2.2) and (2.15), respectively.

Now, by row-block elimination using Z_1 as pivot, we can compute

$$\begin{aligned} Z^{-1} &= \begin{bmatrix} Z_1^{-1}[I - Y(I + XY)^{-1}X] & Z_1^{-1}Y(I + XY)^{-1} \\ -Z_4^{-1}(I + XY)^{-1}X & Z_4^{-1}(I + XY)^{-1} \end{bmatrix} \\ &= \begin{bmatrix} Z_1^{-1}(I + YX)^{-1} & Z_1^{-1}Y(I + XY)^{-1} \\ -Z_4^{-1}(I + XY)^{-1}X & Z_4^{-1}(I + XY)^{-1} \end{bmatrix}. \end{aligned} \tag{2.20}$$

Substituting Z^{-1} in (2.20) into (2.18), we obtain

$$\begin{bmatrix} A_k & 0 \\ -H_k & I \end{bmatrix} = \begin{bmatrix} \Gamma_{1k} & \Gamma_{3k} \\ \Gamma_{2k} & \Gamma_{4k} \end{bmatrix} \begin{bmatrix} J_s^{2^k} Z_1^{-1}(I+YX)^{-1} & J_s^{2^k} Z_1^{-1} Y(I+XY)^{-1} \\ -Z_4^{-1}(I+XY)^{-1}X & Z_4^{-1}(I+XY)^{-1} \end{bmatrix} \tag{2.21}$$

and

$$\begin{bmatrix} I & G_k \\ 0 & A_k^T \end{bmatrix} = \begin{bmatrix} \Gamma_{1k} & \Gamma_{3k} \\ \Gamma_{2k} & \Gamma_{4k} \end{bmatrix} \begin{bmatrix} Z_1^{-1}(I+YX)^{-1} & Z_1^{-1} Y(I+XY)^{-1} \\ -J_s^{2^k} Z_4^{-1}(I+XY)^{-1}X & J_s^{2^k} Z_4^{-1}(I+XY)^{-1} \end{bmatrix}. \tag{2.22}$$

From the (2,1)-block of (2.22), we obtain

$$\Gamma_{2k} Z_1^{-1}(I+YX)^{-1} = \Gamma_{4k} J_s^{2^k} Z_4^{-1} X (I+YX)^{-1} \tag{2.23}$$

implying that

$$\Gamma_{2k} = \Gamma_{4k} J_s^{2^k} Z_4^{-1} X Z_1. \tag{2.24}$$

Consequently, (2.24) and the (2,2)-block of (2.21) lead to

$$\Gamma_{4k}\left[Z_4^{-1}(I+XY)^{-1} + J_s^{2^k} Z_4^{-1} X Z_1 J_s^{2^k} Z_1^{-1} Y (I+XY)^{-1}\right] = I. \tag{2.25}$$

It then follows from (2.24) and (2.25) that

$$\Gamma_{4k} = (I+XY)Z_4 + O\left(\|J_s^{2^{k+1}}\|\right), \quad \Gamma_{2k} = O\left(\|J_s^{2^k}\|\right) \tag{2.26}$$

for sufficiently large k.

Similarly, from the (1,2)-block of (2.21), we have

$$\Gamma_{3k} = -\Gamma_{1k} J_s^{2^k} Z_1^{-1} Y Z_4. \tag{2.27}$$

From the (1,1)-block of (2.22), we obtain

$$\Gamma_{1k}\left[Z_1^{-1}(I+YX)^{-1} + J_s^{2^k} Z_1^{-1} Y Z_4 J_s^{2^k} Z_4^{-1}(I+XY)^{-1} X\right] = I. \tag{2.28}$$

It follows from (2.27) and (2.28) that

$$\Gamma_{1k} = (I+YX)Z_1 + O\left(\|J_s^{2^{k+1}}\|\right), \quad \Gamma_{3k} = O\left(\|J_s^{2^k}\|\right) \tag{2.29}$$

for sufficiently large k. From (2.26) and the (2,1)-block of (2.21), we obtain

$$H_k = \Gamma_{4k} Z_4^{-1} (I + XY)^{-1} X - \Gamma_{2k} J_s^{2^k} Z_1^{-1} (I + YX)^{-1} = X + O\left(\| J_s^{2^{k+1}} \| \right) \tag{2.30}$$

for sufficiently large k. Equation (2.29) and the (1,2)-block of (2.22) then lead to

$$G_k = \Gamma_{1k} Z_1^{-1} (I + YX)^{-1} Y + \Gamma_{3k} J_s^{2^k} Z_4^{-1} (I + XY)^{-1} = Y + O\left(\| J_s^{2^{k+1}} \| \right) \tag{2.31}$$

for k sufficiently large.

Finally, (2.29) and the (1,1)-block of (2.21) imply

$$A_k = \Gamma_{1k} J_s^{2^k} Z_1^{-1} (I + YX)^{-1} - \Gamma_{3k} Z_4^{-1} (I + XY)^{-1} X = (I + YX) Z_1 J_s^{2^k} Z_1^{-1} = O\left(\| J_s^{2^k} \| \right). \tag{2.32}$$

Since the spectral radius of J_s equals to $|\lambda_n| < 1$, (2.30)–(2.32) imply the results in (i)–(iii), as well as the $O\left(|\lambda_n|^{2^k} \right)$ rate of convergence. □

The following Lemma proves that the stabilizability and detectability properties are preserved by the SDA throughout its iterative process.

Lemma 2.3. *The stabilizability of (A, B) implies that $(\widehat{A}, \widehat{B})$ is stabilizable, where $\widehat{G} = \widehat{B}\widehat{B}^T \geq 0$ is a FRD of $\widehat{G}$. The detectability of (A, C) implies that $(\widehat{A}, \widehat{C})$ is detectable, where $\widehat{H} = \widehat{C}^T\widehat{C} \geq 0$ is a FRD of $\widehat{H}$.*

Proof. See Appendix. □

Theorem 2.4. *Let $M = \begin{bmatrix} A & 0 \\ -H & I \end{bmatrix}$ and $L = \begin{bmatrix} I & G \\ 0 & A^T \end{bmatrix}$ with $G = BR^{-1}B^T \geq 0$ and $H = C^TC \geq 0$. Assume that (A, B) is stabilizable and (A, C) is detectable. Then the sequences $\{A_k, H_k, G_k\}$ computed by the SDA satisfy (i), (ii), (iii) as in Theorem 2.2.*

Proof. It is well-known that these reasonable assumptions implies that $M - \lambda L$ has no eigenvalues on the unit circle, and that Z_1 and Z_4 are invertible (see, for example, [84, 95], for details). Thus the conditions in Theorem 2.2 are satisfied. □

Remark. Theoretically, the convergence behavior for the SDA and the algorithms in [7, 18, 20] are similar. Nevertheless, Theorem 2.2 directly proves, under the assumptions that $M - \lambda L$ have no unit modulo eigenvalues and Z_1, Z_4 are invertible, that the sequences $\{A_k, H_k, G_k\}$ generated by the SDA converge to zero and the unique s.p.s.d. solutions of the DAREs in (2.2) and (2.15), respectively. Lemma 2.3 shows the preservation of stabilizability and detectability of the iterates (A_k, G_k, H_k) generated by the SDA. Furthermore, in Theorem 2.4, we see that the assumptions in Theorem 2.3 are weaker than the conditions (S) and (D). This distinction of preserving the symplectic structure in SSF, as well as the difference in operation counts, are responsible for the superior performance of the SDA.

Computation of $\widehat{A}$, $\widehat{G}$ and $\widehat{H}$

We now propose a structured and efficient procedure for the computation of $\widehat{A}$, $\widehat{G}$ and $\widehat{H}$ in (2.9)–(2.11), respectively, where $G = BB^T \geq 0$, $H = C^TC \geq 0$ are FRDs. Let $W = (I + GH)^{-1}$. It is easily seen that $HW = W^TH$ and $GW^T = WG$ are s.p.s.d.. By the SMWF we can derive the formulae

$$W = (I + GH)^{-1} = I - B(I + B^THB)^{-1}B^TH, \tag{2.33}$$

$$GW^T = G - GC^T(I + CGC^T)^{-1}CG = B(I + B^THB)^{-1}B^T, \tag{2.34}$$

$$W^TH = H - HB(I + B^THB)^{-1}B^TH = C^T(I + CGC^T)^{-1}C. \tag{2.35}$$

When B and C start with low ranks, we can improve the efficiency of our computation further by the following compression process. Compute the Cholesky decomposition of $W_G \equiv (I + B^THB) = K_B^TK_B$ and $W_H \equiv (I + CGC^T) = K_CK_C^T$. Apply (2.33)–(2.35) to (2.9)–(2.11), we compute

$$\widehat{A} = A^2 - AB(I + B^THB)^{-1}B^THA, \tag{2.36}$$

$$\begin{aligned}\widehat{G} &= G + AB(I + B^THB)^{-1}B^TA^T \\ &= \begin{bmatrix} B, & ABK_B^{-1} \end{bmatrix} \begin{bmatrix} B^T \\ K_B^{-T}B^TA^T \end{bmatrix} \equiv \widehat{B}\widehat{B}^T \geq 0 \quad \text{(FRD)}\end{aligned} \tag{2.37}$$

and

$$\begin{aligned}\widehat{H} &= H + A^T C^T (I + CGC^T)^{-1} CA \\ &= \left[\begin{array}{cc} C^T, & A^T C^T K_C^{-T} \end{array}\right] \left[\begin{array}{c} C \\ K_C^{-1} CA \end{array}\right] \equiv \widehat{C}^T \widehat{C} \geq 0 \quad \text{(FRD)},\end{aligned} \tag{2.38}$$

where $\widehat{B}$ and $\widehat{C}^T$ are the full column rank compressions of matrices $\left[\begin{array}{cc} B, & ABK_B^{-1} \end{array}\right]$ and $\left[\begin{array}{cc} C^T, & A^T C^T K_C^{-T} \end{array}\right]$, respectively. In general, $\text{rank}(\widehat{B}) > \text{rank}(B)$ and $\text{rank}(\widehat{C}) > \text{rank}(C)$, and the compression process becomes unprofitable when the ranks of $\widehat{B}$ and $\widehat{C}$ reach n.

Remark. From (2.34) and (2.35), it is necessary to compute the Choleskey decompositions of the symmetric positive definite matrices W_G and W_H when updating G and H. This requires $m^3/3 + r^3/3$ flops. If, instead of Cholesky factors, we compute the square roots of W_G and W_H in (2.34) and (2.35), then an additional $12(m^3 + r^3)$ flops are required. Here the square roots of W_G and W_H are obtained from

$$\begin{aligned} W_G &= K_B^T K_B = V_B \Sigma_B V_B^T = (V_B \Sigma_B V_B^T)(V_B \Sigma_B V_B^T) = X_B^2 \\ W_H &= K_C K_C^T = U_C \Sigma_C U_C^T = (U_C \Sigma_C U_C^T)(U_C \Sigma_C U_C^T) = X_C^2, \end{aligned}$$

and the SVDs $K_B = U_B \Sigma_B V_B^T$ and $K_C = U_C \Sigma_C V_C^T$. Cheaper methods for calculating the square roots are available but the corresponding implications on numerical stability and cost benefits are questionable. As a result, we do not choose the square root alternative in our algorithm.

Error Analysis of SDA

We now consider the errors in calculating $\widehat{A}$, $\widehat{G}$ and $\widehat{H}$ as in (2.9)–(2.11), respectively. From (2.9)–(2.11) we see that the matrix $W \equiv (I + GH)^{-1}$ occurs frequently in the SDA algorithm, for some generic s.p.s.d. matrices G and H. From (2.33)–(2.35), instead of inverting the nonsymmetric $(I + GH)$, we can invert the s.p.d. matrices $(I + B^T HB)$ and

$(I + CGC^T)$, when updating of $\widehat{A}$, $\widehat{G}$ and $\widehat{H}$. The conditioning of $(I + B^T HB)$ in (2.33) and (2.34) (or $(I + CGC^T)$ in (2.35)) is well-known, with the condition number being

$$\kappa = \frac{1 + \sigma^2_{\max}(CB)}{1 + \sigma^2_{\min}(CB)} \tag{2.39}$$

and σ denoting the singular values. The error analysis in the updating (A, G, H) to $(\widehat{A}, \widehat{G}, \widehat{H})$ in (2.9)–(2.11) is thus reduced to the routine discussion about the accumulation of errors in forming sums and products. With δ indicating errors, $|F|$ denoting the matrix with all signs of elements in F ignored and Δ denoting the maximum error in the starting data A, G and H, we typically have the asymptotic inequalities

$$|||\delta\widehat{A}|||, |||\delta\widehat{G}|||, |||\delta\widehat{H}||| \preceq (c_1 + \kappa c_2)\Delta$$

with c_1 and c_2 being polynomials in n of low degrees. Note that the coefficients c_1 and c_2 are dependent on the sizes of A, G and H.

When the condition number κ in (2.39) is bounded by an acceptable number, the accumulation of error will be dampened by the fast rate of convergence at the final stage of the iterative process. Danger, if any, lies in the early stage of the process before the $\lambda_n^{2^k}$ convergence factor dominates. It is unlikely to have any ill-effect, as the accumulated error in the matrix additions and multiplications should be of magnitude around a small multiple of the machine accuracy.

As the SSF properties are preserved in the SDA, any error will be a structured one, only pushing the iteration towards a solution of a neighboring SSF system. Thus the algorithm is stable in this sense, when the errors are not too large and when stabilizability and detectability are maintained. For large ks, as $A_k \to 0$, G_k and H_k converge to the unique s.p.s.d. solutions of (2.15) and (2.2), respectively. Danger again will only comes at the initial stage of the iteration. Corresponding checks may be prudent in the algorithm.

Operation Counts

The matrix disk function method in [18, 20] is developed to solve the DARE (2.2) by using a swapping technique built on the QR decomposition. We refer the algorithm presented in

[18, 20] as QR-SWAP. We shall perform a flop-count for the SDA as well as the QR-SWAP algorithm. For the counts for components like LU- and QR-decompositions, consult [59] for details. For the SDA, we have the following count for one iteration:

Calculation in SDA	Flops
GH	n^3
LU decomposition of $I+GH$	$\frac{2}{3}n^3$
$(I+HG)^{-1}A^T$	n^3
$\widehat{A}=A(I+GH)^{-1}A$	n^3
$(I+GH)^{-1}A$	n^3
AG	n^3
$\widehat{G}=G+AG(I+HG)^{-1}A^T$	$\frac{1}{2}n^3$
HA	n^3
$\widehat{H}=H+A^T(I+HG)^{-1}HA$	$\frac{1}{2}n^3$
The total count =	$\frac{23}{3}n^3$

There is a small saving by (2.33)–(2.35) at the early stage of the iteration, when G and H have low ranks. We ignore this saving in the above count. Note that the symmetry in $\widehat{G}$ and $\widehat{H}$ saves n^3 flops. We have also ignored any $O(n^2)$ operation counts and the memory counts.

For the QR-SWAP algorithm [18, 20], we have the following count for one iteration:

Calculation in QR-SWAP	Flops
$Q \begin{bmatrix} L \\ -M \end{bmatrix} = \begin{bmatrix} R \\ 0 \end{bmatrix}$	$\frac{80}{3}n^3$
Forming $Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$	$\frac{224}{3}n^3$
$Q_{21}L$	$8n^3$
$Q_{22}M$	$8n^3$
The total count =	$\frac{352}{3}n^3$

There is some saving for the QR-SWAP algorithm in the first iteration, making use of the structure in M and L. This structure is lost in the later stages. There is also some saving in the accumulation of Householder factors when forming Q, as only part of Q is later required. This accounts for part of the over-estimation in the table above, as compared to the operation count of $\frac{320}{3}n^3$ flops from [18, 20].

The operation count for the SDA is about 7% of that for QR-SWAP. This is mainly due to the fact that the main steps in QR-SWAP involve the QR decomposition of $\begin{bmatrix} L \\ -M \end{bmatrix} \in \mathcal{R}^{4n\times 2n}$ and the formation of $Q \in \mathcal{R}^{4n\times 4n}$, all in higher dimensions. The operations in the SDA are all within $\mathcal{R}^{n\times n}$.

It is difficult to conduct an operation count for `dare`, mainly because of the iterative nature of the Schur decomposition before invariant subspaces and solutions can be obtained. Operation counts per iteration should be of the same order as QR-SWAP. Peripheral operations in MATLAB itself also add heavily to the count and making a detailed comparison difficult.

3 Swap and Collapse

Recall that the s.p.s.d. solutions $\{X_j\}_{j=1}^p$ to the P-DAREs (1.5) can be obtained from the invariant subspace associated with the stable eigenvalues for Π_j in (1.9) when all A_j are nonsingular. In general, the representation of Π_j in (1.9) can also be applied when some of A_j are singular or even not squared as in (1.4). The swap and collapse process [19, 20], which does not form the product Π_j in (1.9) explicitly, can be used to compute X_{j-1} by swapping the order of the products and collapsing them into a single symplectic matrix pair $(\widehat{M}_j, \widehat{L}_j)$ such that $\Pi_j = \widehat{L}_j^{-1}\widehat{M}_j$, where $\widehat{L}_j$, $\widehat{M}_j \in \mathcal{R}^{2n_{j-1}\times 2n_{j-1}}$. The process relies on the following Lemma [20, Lemma 1]:

Lemma 3.1. *Consider $E \in \mathcal{R}^{s\times q}$, $F \in \mathcal{R}^{t\times q}$ and let*

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} E \\ -F \end{bmatrix} = \begin{bmatrix} R \\ 0 \end{bmatrix}$$

be a QR factorization of $[E^T, -F^T]^T$ where $R \in \mathcal{R}^{q\times q}$, $Q_{11} \in \mathcal{R}^{q\times s}$, $Q_{12} \in \mathcal{R}^{q\times t}$, $Q_{21} \in \mathcal{R}^{(s+t-q)\times s}$ and $Q_{22} \in \mathcal{R}^{(s+t-q)\times t}$. Then

$$Q_{22}^{-1}Q_{21} = FE^{-1}. \tag{3.1}$$

Here, the inverse of E or Q_{22} is purely notational. In fact, the relation in (3.1) denotes the relation $Q_{21}E = Q_{22}F$. We use the relation $Q_{22}^{-1}Q_{21} = FE^{-1}$ in the swap and collapse process even when E or Q_{22} are singular or not squared. Using Lemma 3.1, the order of the products in (1.9) can be swapped, with all Ls collapsed together to form $\widehat{L}_j^{-1}$. As an illustration, let us consider the following Π_1 for a period $p = 4$:

$$\Pi_1 = L_4^{-1}M_4L_3^{-1}M_3L_2^{-1}M_2L_1^{-1}M_1. \tag{3.2}$$

Note that the sizes of matrices M_j and L_j are given in (1.6) with $n_j = n_{j+4}$. Applying Lemma 3.1, we can swap the order in the product $M_2L_1^{-1} = (L_1^{(1)})^{-1}M_2^{(1)}$ to obtain

$$\Pi_1 = L_4^{-1}M_4L_3^{-1}M_3L_2^{-1}(L_1^{(1)})^{-1}M_2^{(1)}M_1.$$

Collapsing $L_1^{(1)}L_2$ and $M_2^{(1)}M_1$ into, respectively, $L_{1:2}$ and $M_{1:2}$, we obtain

$$\Pi_1 = L_4^{-1}M_4L_3^{-1}M_3L_{1:2}^{-1}M_{1:2}.$$

Repeat the process, swapping $M_3L_{1:2}^{-1} = (L_{1:2}^{(1)})^{-1}M_3^{(1)}$ and then collapsing the resulting terms, we obtain, with $L_{1:3}^{-1}M_{1:3} = L_3^{-1}(L_{1:2}^{(1)})^{-1}M_3^{(1)}M_{1:2}$,

$$\Pi_1 = L_4^{-1}M_4L_3^{-1}(L_{1:2}^{(1)})^{-1}M_3^{(1)}M_{1:2} = L_4^{-1}M_4L_{1:3}^{-1}M_{1:3}.$$

A final swap for $M_4L_{1:3}^{-1} = (L_{1:3}^{(1)})^{-1}M_4^{(1)}$ and the associated collapse step will produce

$$\Pi_1 = L_4^{-1}(L_{1:3}^{(1)})^{-1}M_4^{(1)}M_{1:3} = L_{1:4}^{-1}M_{1:4},$$

where $L_{1:4}$, $M_{1:4} \in \mathcal{R}^{2n_4\times 2n_4}$. The solution X_4 can then be calculated via the stable invariant subspace of $(M_{1:4}, L_{1:4})$, with other X_k $(k \neq 1)$ obtained from (1.1).

As indicated in [19, 20], notice that the swap and collapse step can be performed for different products in Π_j in parallel. For example in (3.2), we can swap and collapse $M_4L_3^{-1}$ and $M_2L_1^{-1}$ simultaneously to obtain

$$\begin{aligned}\Pi_1 &= L_4^{-1}M_4L_3^{-1}M_3L_2^{-1}M_2L_1^{-1}M_1 = L_4^{-1}(L_3^{(1)})^{-1}M_4^{(1)}M_3L_2^{-1}(L_1^{(1)})^{-1}M_2^{(1)}M_1\\ &= L_{3:4}^{-1}M_{3:4}L_{1:2}^{-1}M_{1:2}.\end{aligned}$$

A final swap and collapse associated with $M_{3:4}L_{1:2}^{-1}$ then produces

$$\Pi_1 = L_{3:4}^{-1}(L_{1:2}^{(1)})^{-1}M_{3:4}^{(1)}M_{1:2} = L_{1:4}^{-1}M_{1:4}.$$

More importantly, notice that the QR factorization in Lemma 3.1 can be replaced by other factorizations. We now develop a structure-preserving procedure, which is closely related to the QR-SWAP algorithms [19, 20], based on the LU-like factorization as in (2.3)–(2.5) to reduce the periodic symplectic matrix pairs $\{(M_j, L_j)\}_{j=1}^p$ in (1.6) to a single symplectic matrix pair $(M_{1:p}, L_{1:p}) \in \mathcal{R}^{2n_p\times 2n_p} \times \mathcal{R}^{2n_p\times 2n_p}$ in SSF.

Given M_1, L_1, M_2 and L_2 as in (1.6), we have the factorization

$$T\begin{bmatrix} L_1 \\ -M_2 \end{bmatrix} = \begin{bmatrix} R \\ 0 \end{bmatrix}$$

or

$$\left[\begin{array}{cc|cc} I_{n_1} & 0 & 0 & 0 \\ -H_2 & 0 & 0 & I_{n_1} \\ \hline A_2(I_{n_1}+G_1H_2)^{-1} & 0 & I_{n_2} & A_2G_1(I_{n_1}+H_2G_1)^{-1} \\ -A_1^T(I_{n_1}+H_2G_1)^{-1}H_2 & I_{n_4} & 0 & A_1^T(I_{n_1}+H_2G_1)^{-1} \end{array}\right] \left[\begin{array}{cc} I_{n_1} & G_1 \\ 0 & A_1^T \\ \hline -A_2 & 0 \\ H_2 & -I_{n_1} \end{array}\right]$$

$$= \left[\begin{array}{cc} I_{n_1} & G_1 \\ 0 & -(I_{n_1}+H_2G_1) \\ \hline 0 & 0 \\ 0 & 0 \end{array}\right]. \tag{3.3}$$

Again, similar to (2.5), the transformation T represents row operations on $\left[\begin{array}{c} L_1 \\ -M_2 \end{array}\right]$, and the factorization (3.3) can easily be checked by direct multiplication with the help of the SMWF.

Similar to Lemma 3.1, it is then obvious, after the swap, that we have

$$M_2L_1^{-1} = (L_1^{(1)})^{-1}M_2^{(1)} = Q_{22}^{-1}Q_{21}$$

with $Q_{22} = L_1^{(1)}$ being the bottom-right $(n_2+n_4)\times(n_1+n_2)$ block in T of (3.3), $Q_{21} = M_2^{(1)}$ being the bottom-left $(n_2+n_4)\times(n_1+n_4)$ block in T, and

$$L_1^{(1)} = \left[\begin{array}{cc} I_{n_2} & A_2G_1(I_{n_1}+H_2G_1)^{-1} \\ 0 & A_1^T(I_{n_1}+H_2G_1)^{-1} \end{array}\right], \quad M_2^{(1)} = \left[\begin{array}{cc} A_2(I_{n_1}+G_1H_2)^{-1} & 0 \\ -A_1^T(I_{n_1}+H_2G_1)^{-1}H_2 & I_{n_4} \end{array}\right]. \tag{3.4}$$

Consequently, we obtain

$$L_{1:2} := L_1^{(1)}L_2 = \left[\begin{array}{cc} I_{n_2} & G_2+A_2G_1(I_{n_1}+H_2G_1)^{-1}A_2^T \\ 0 & A_1^T(I_{n_1}+H_2G_1)^{-1}A_2^T \end{array}\right] \equiv \left[\begin{array}{cc} I_{n_2} & \widehat{G}_2 \\ 0 & \widehat{A}_2^T \end{array}\right] \tag{3.5}$$

and

$$M_{1:2} := M_2^{(1)}M_1 = \left[\begin{array}{cc} A_2(I_{n_1}+G_1H_2)^{-1}A_1 & 0 \\ -[H_1+A_1^T(I_{n_1}+H_2G_1)^{-1}H_2A_1] & I_{n_4} \end{array}\right] \equiv \left[\begin{array}{cc} \widehat{A}_2 & 0 \\ -\widehat{H}_2 & I_{n_4} \end{array}\right]. \tag{3.6}$$

Notice that $(M_{1:2}, L_{1:2})$ is again in SSF and satisfies

$$L_{1:2}^{-1} M_{1:2} = L_2^{-1} M_2 L_1^{-1} M_1 \ . \tag{3.7}$$

In (3.4)–(3.6), we have performed the transformation

$$\left[\begin{array}{c|cccccc} -M_1 & L_1 & & & & & \\ & -M_2 & L_2 & & & & \\ & & -M_3 & L_3 & & & \\ & & & \ddots & \ddots & & \\ & & & & -M_p & L_p \end{array}\right] \xrightarrow[T]{} \left[\begin{array}{c|cccccc} * & * & * & & & & \\ -M_{1:2} & 0 & L_{1:2} & & & & \\ & & -M_3 & L_3 & & & \\ & & & \ddots & \ddots & & \\ & & & & -M_p & L_p \end{array}\right] \tag{3.8}$$

using row operations in T, with $L_{1:2}$ and $M_{1:2}$ given in (3.5) and (3.6).

The calculation in the structure-preserving swap and collapse algorithm (SSCA), continuing the swaps and collapses in (3.4)–(3.6), can then be summarized as:

For $j = 2, 3, \cdots, p$,

$$\widehat{A}_j = A_j (I_{n_{j-1}} + \widehat{G}_{j-1} H_j)^{-1} \widehat{A}_{j-1} \in \mathcal{R}^{n_j \times n_p}, \tag{3.9}$$

$$\widehat{G}_j = G_j + A_j \widehat{G}_{j-1} (I_{n_{j-1}} + H_j \widehat{G}_{j-1})^{-1} A_j^T \in \mathcal{R}^{n_j \times n_j}, \tag{3.10}$$

$$\widehat{H}_j = \widehat{H}_{j-1} + \widehat{A}_{j-1}^T (I_{n_{j-1}} + H_j \widehat{G}_{j-1})^{-1} H_j \widehat{A}_{j-1} \in \mathcal{R}^{n_p \times n_p} \tag{3.11}$$

with $\widehat{A}_1 = A_1$, $\widehat{G}_1 = G_1$ and $\widehat{H}_1 = H_1$, and the \hat{·} denoting the result of a swap and collapse step.

Finally, the following SSCA reduces the periodic symplectic matrix pairs $\{(M_j,\ L_j)\}_{j=1}^p$ as in (1.6) into a single symplectic matrix pair in SSF

$$(\widehat{M},\ \widehat{L}) \equiv (M_{1:p},\ L_{1:p}) = \left(\begin{bmatrix} \widehat{A}_p & 0 \\ -\widehat{H}_p & I_{n_p} \end{bmatrix},\ \begin{bmatrix} I_{n_p} & \widehat{G}_p \\ 0 & \widehat{A}_p^T \end{bmatrix} \right).$$

SSCA Algorithm

Input : A_j; G_j, $H_j \geq 0$, $j = 1, \cdots, p$;

Output : $\widehat{A}_p$; $\widehat{G}_p$, $\widehat{H}_p \geq 0$;

Initialize $\widehat{A}_1 \leftarrow A_1$, $\widehat{G}_1 \leftarrow G_1$, $\widehat{H}_1 \leftarrow H_1$;

For $j = 2, 3, \cdots, p$

Compute $W \leftarrow I_{n_{j-1}} + H_j \widehat{G}_{j-1}$;

Solve $WV_1 = A_j^T$, $WV_2 = H_j$ for V_1, V_2;

$\widehat{A}_j \leftarrow V_1^T \widehat{A}_{j-1}$, $\widehat{G}_j \leftarrow G_j + A_j \widehat{G}_{j-1} V_1$, $\widehat{H}_j \leftarrow \widehat{H}_{j-1} + \widehat{A}_{j-1}^T V_2 \widehat{A}_{j-1}$;

End

End of SSCA Algorithm

Remarks. (i) It is vital to preserve the SSF property of the symplectic matrix pairs in the SSCA, by maintaining the symmetry of $\widehat{G}_j$ and $\widehat{H}_j$ for $j = 2, 3, \ldots, p$. For solving P-DAREs, applying the SDA to the collapsed system produces $X_p = X_0$, from which the other X_j ($j = p-1, \ldots, 1$) can be found through (1.1). We call this combination as SSCA+SDA.

(ii) It can be observed that the operations in the SSCA are closely related to those in the SDA in Section 2. (The iterative process in the SDA is like the swapping and collapsing in the SSCA, with periodicity $p = 1$.) The error analysis of the SDA in the previous Section can be shared with that for the SSCA.

(iii) In Lemma 3.1, an orthogonal reduction technique, based on the QR-factorization, is used to swap and collapse two matrix pairs. However, the process is not backward stable because only the lower half part of the orthogonal transformation is involved. Consequently, the swap and collapse procedure of the QR-SWAP algorithm [19, 20] is not backward stable and the final collapsed matrix pair $(M_{1:p}, L_{1:p})$ is generally not in SSF.

On the other hand, the swapping and collapsing of two matrix pairs using GSVD has been proven to be numerically backward stable [19]. However, the computational

cost of the GSVD is much higher than that of the QR-factorization. Also, the question of stability for the collapsing of products of more than two matrix pairs is still open. Thus, the swapping and collapsing procedure using GSVD is not considered here.

In the SSCA, non-orthogonal transformations are used and matrix products are computed in each step. However, the nice structures in the standard symplectic form (SSF) are preserved in each step. Furthermore, the SSCA only involves inverses of s.p.d. matrices, with little error accumulation. For solving P-DAREs, applying the SDA to the collapsed system produces $X_p = X_0$, from which the other X_j $(j = p-1, \ldots, 1)$ can be found through (1.1). Accumulation of error for moderate values of p should be acceptable. A good check of accuracy will be to calculate $\widetilde{X}_p$ again by substituting X_1 into (1.1) and compare that with the X_p from the SDA.

(iv) The SSCA is utilized a Gaussian-like decomposition to perform the swaps in (3.3), in contrast with the QR decompositions used in [18, 19, 20]. Although important differences in structure-preserving set the algorithms apart in performance, they share the same parallelism. Swaps and collapses can be carried out at different point in (3.8) simultaneously (see [20] for details of parallelism, and the related remarks in Section 6). This parallelism is obvious in the SSCA.

Preservation of Positivity, Stabilizability and Detectability

The following Lemma proves that the important stabilizability and detectability are preserved by the SSCA.

Lemma 3.2. *The (P-S) property of the* $\{(A_j, B_j)\}_{j=1}^{p}$ *implies that* $(\widehat{A}_p, \widehat{B}_p)$ *is stabilizable, where* $\widehat{G}_p = \widehat{B}_p \widehat{B}_p^T \geq 0$ *is the FRD of* $\widehat{G}_p$*. The (P-D) property of the* $\{(A_j, C_j)\}_{j=1}^{p}$ *implies that* $(\widehat{A}_p, \widehat{C}_p)$ *is detectable, where* $\widehat{H}_p = \widehat{C}_p^T \widehat{C}_p \geq 0$ *is the FRD of* $\widehat{H}_p$ *.*

Proof. See Appendix. □

4 Numerical Experiments for DAREs

The aim of this Section is to illustrate the superior performance of the SDA algorithm, as compared to the QR-SWAP algorithm [20] and the MATLAB control toolbox command `dare` [88]. The solution of a small number of difficult DAREs, some from the benchmark set of problems in [21], are considered. Some problems have parameters which control their degree of difficulty and conditioning. Tables of residuals, relative errors and iteration numbers are presented for selected values of the parameter. For examples with varying dimensions, graphs of accuracies, CPU-times and efficiency ratios against the problem size are also presented.

Numerical results confirm the accuracy and efficiency of the SDA as predicted in Section 2. In particular, the SDA is up to ten times more efficient than the QR-SWAP, as predicted by the flop-counts in Section 2. The SDA also seems to be more efficient than `dare` but the comparison with part of a general purpose package cannot be done on a equal footing. All in all, the SDA is efficient without equal comparing to other methods, for the difficult problems we have tested. We see no reason why it should be performing differently for other problems in general, especially after other refinements, such as the possibilities in parallel computing discussed in Section 6, are implemented.

Now some details in the numerical experiments are listed. When the exact solution, denoted by X, is known, the relative error of an approximate solution $\widetilde{X}$ is calculated by

$$\text{Rel. err.} \equiv \frac{\|\widetilde{X} - X\|_F}{\|X\|_F} .$$

The associated residual is calculated by

$$\text{Residual} \equiv \|A^T \widetilde{X}(I + G\widetilde{X})^{-1}A + H - \widetilde{X}\|_F .$$

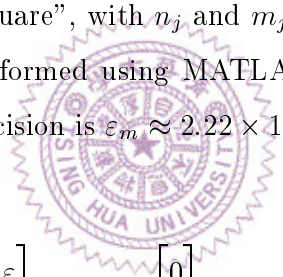
We try our best to compare the CPU time used by different methods for approximate solutions of similar accuracies. Often, it is impossible to match these accuracies and we have been forced to compare more accurate results from the SDA with less accurate ones from other methods. The opposite situation of comparing less accurate results from the

SDA seldom occurs. This issue of relative accuracies in the comparison for DAREs is less severe than for P-DAREs in Section 5.

For the Tables in the following examples, data for various methods are lists in columns with obvious headings. The heading "`dare`" is for the `dare` command in MATLAB [88], "QR" is for the QR-SWAP method in [20], and "SDA" stands for the SDA algorithm. There is no iteration numbers to report for `dare` and an '$*$' in the Tables indicates a failure of convergence. Failures occur frequently for `dare` for difficult problems. In the graphs, "ratio_dare" is the ratio between the CPU-times for `dare` and the SDA, and "ratio_QR" is defined similarly. Notice the logarithmic scales used in these graphs.

Here we only reported some representative numerical examples and retained the numbering of examples in [46], where more numerical examples are presented. In order to demonstrate the behavior of quadratic convergence for the SDA and QR-SWAP algorithms, we display the relative errors (d_j^{SDA} and d_j^{QR} in Frobenius norm) in two examples. Note that all examples are "square", with n_j and m_j being invariant of j.

All computations were performed using MATLAB/Version 6.0 on a Compaq/DS20 workstation. The machine precision is $\varepsilon_m \approx 2.22 \times 10^{-16}$.



Example 4.2. Let

$$A = \begin{bmatrix} 0 & \varepsilon \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad R = 1, \quad H = I_2.$$

The stabilizing solution is given by

$$X = \begin{bmatrix} 1 & 0 \\ 0 & 1+\varepsilon^2 \end{bmatrix}$$

and the closed-loop spectrum is $\{0, 0\}$. For $\varepsilon = 100$, this is Example 2 from [60]. As $\varepsilon \to \infty$, this becomes an example of a DARE which is badly scaled in the sense of [99], due to the fact that $\|A\|_F \gg \|G\|_F, \|H\|_F$. The numerical results with $\varepsilon = 100, 10^4, 10^6$ are given in Table 1. For $\varepsilon = 100$, the behavior of quadratic convergence for the QR-SWAP algorithm and SDA is shown in Table 2.

		`dare`	QR	SDA
$\varepsilon = 100$	Residual	6.72×10^{-9}	3.25×10^{-17}	0.00×10^{0}
	Rel. err.	6.72×10^{-13}	3.25×10^{-21}	0.00×10^{0}
	Iter. no.	-	4	2
$\varepsilon = 10^4$	Residual	4.40×10^{-1}	2.98×10^{-8}	0.00×10^{0}
	Rel. err.	4.40×10^{-9}	2.98×10^{-16}	0.00×10^{0}
	Iter. no.	-	4	2
$\varepsilon = 10^6$	Residual	6.11×10^{7}	1.22×10^{-4}	0.00×10^{0}
	Rel. err.	6.11×10^{-5}	1.22×10^{-16}	0.00×10^{0}
	Iter. no.	-	4	2

Table 1: Results for Example 4.2.

Example 4.3. The following example is identical to Example 13 of [21] which was presented originally in [99]. Let

$$A_0 = \mathrm{diag}(0, 1, 3), \quad V = I - \frac{2}{3}vv^T, \quad v^T = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

Then

$$A = VA_0V, \quad G = \frac{1}{\varepsilon}I_3, \quad H = \varepsilon I_3.$$

The factorization $H = C^TC$ with $C = \sqrt{\varepsilon}V$, and similarly, $G = BR^{-1}B^T$ with $B = I_3$

j	d_j^{QR}	d_j^{SDA}
1	1.00×10^{-2}	1.00×10^{0}
2	1.41×10^{-2}	0.00×10^{0}
3	1.42×10^{-18}	0.00×10^{0}
4	$\leq \varepsilon_m$	0.00×10^{0}

Table 2: Results for Example 4.2 with $\varepsilon = 100$.

and $R = \varepsilon I_3$. The exact solution is given by

$$X = V \ \mathrm{diag}(x_1, x_2, x_3)\ V$$

where

$$x_1 = \varepsilon, \quad x_2 = \varepsilon\frac{(1+\sqrt{5})}{2}, \quad x_3 = \varepsilon\frac{(9+\sqrt{85})}{2}.$$

The numerical results with $\varepsilon = 1.0, 10^4, 10^6$ are given in Table 3. For $\varepsilon = 1.0$, the behavior of quadratic convergence for the QR-SWAP algorithm and SDA is shown in Table 4.

		`dare`	QR	SDA
$\varepsilon = 1.0$	Residual	2.57×10^{-15}	3.09×10^{-15}	2.23×10^{-15}
	Rel. err.	2.01×10^{-16}	4.04×10^{-16}	1.86×10^{-16}
	Iter. no.	-	7	6
$\varepsilon = 10^4$	Residual	2.18×10^{-11}	1.79×10^{-3}	1.93×10^{-11}
	Rel. err.	2.39×10^{-16}	2.11×10^{-8}	1.72×10^{-16}
	Iter. no.	-	7	6
$\varepsilon = 10^6$	Residual	2.66×10^{-9}	1.22×10^{3}	1.47×10^{-9}
	Rel. err.	2.80×10^{-16}	1.42×10^{-4}	1.64×10^{-16}
	Iter. no.	-	6	6

Table 3: Results for Example 4.3.

Example 4.5. The following example is identical to Example 15 of [21] which was presented originally in [96, Example 3]. Consider the DARE defined by

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix} \in \mathcal{R}^{n\times n}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad R = r, \quad H = I_n.$$

j	d_j^{QR}	d_j^{SDA}
1	6.07×10^{-1}	4.87×10^{-1}
2	5.41×10^{-2}	3.83×10^{-1}
3	7.02×10^{-3}	4.93×10^{-3}
4	1.41×10^{-4}	3.24×10^{-7}
5	6.52×10^{-8}	2.65×10^{-14}
6	1.35×10^{-14}	0.00×10^{0}
7	$\leq \varepsilon_m$	0.00×10^{0}

Table 4: Results for Example 4.3 with $\varepsilon = 1.0$.

The stabilizing solution has a very simple form, namely,

$$X = \mathrm{diag}(1, 2, \cdots, n).$$

Note that the choice of r does not influence the stabilizing solution X but for $r < 1$, the condition number of DARE behaves like $1/r$. In Figure 1, we report the comparison of CPU times and its ratio with respect to the SDA for $n = 50, 100, 150, 200, 250, 300$. We also list the residuals (res) and relative errors (RE) in Table 5. Note that the residuals and relative errors for the SDA are in machine accuracy.

For smaller parameter r, say $r = 10^{-12}$, the report of CPU times, residuals and relative errors are given in the Figure 2 and Table 6, respectively. Again, the residuals and relative errors for the SDA are in machine accuracy.

Example 4.6. In this example we consider a linear system (A, B, C) such that the corresponding symplectic matrix pair (M, L) has a pair of eigenvalues close nearly to the unit circle in the complex plane. In the following the system matrices are constructed step by step via some symplectic structure-preserving equivalence transformations.

n	res_dare	res_QR	res_SDA	RE_dare	RE_QR	RE_SDA
50	3.50×10^{-12}	5.11×10^{-12}	0.00×10^{0}	4.99×10^{-14}	4.33×10^{-14}	0.00×10^{0}
100	3.55×10^{-11}	7.17×10^{-11}	0.00×10^{0}	2.46×10^{-13}	1.80×10^{-13}	0.00×10^{0}
150	1.04×10^{-10}	2.32×10^{-10}	0.00×10^{0}	1.04×10^{-12}	5.06×10^{-13}	0.00×10^{0}
200	2.58×10^{-10}	5.67×10^{-10}	0.00×10^{0}	2.66×10^{-12}	8.41×10^{-13}	0.00×10^{0}
250	5.31×10^{-10}	1.39×10^{-9}	0.00×10^{0}	4.91×10^{-12}	9.08×10^{-13}	0.00×10^{0}
300	1.03×10^{-9}	2.78×10^{-9}	0.00×10^{0}	1.03×10^{-11}	1.67×10^{-12}	0.00×10^{0}

Table 5: Results for Example 4.5 with $r = 1$.

n	res_dare	res_QR	res_SDA	RE_dare	RE_QR	RE_SDA
50	1.20×10^{-11}	5.12×10^{-12}	0.00×10^{0}	2.25×10^{-13}	4.36×10^{-14}	0.00×10^{0}
100	8.13×10^{-11}	6.34×10^{-11}	0.00×10^{0}	4.93×10^{-13}	1.82×10^{-13}	0.00×10^{0}
150	2.17×10^{-10}	2.50×10^{-10}	0.00×10^{0}	2.67×10^{-12}	5.43×10^{-13}	0.00×10^{0}
200	5.68×10^{-10}	6.34×10^{-10}	0.00×10^{0}	4.88×10^{-12}	8.53×10^{-13}	0.00×10^{0}
250	1.32×10^{-9}	1.75×10^{-9}	0.00×10^{0}	1.33×10^{-11}	8.99×10^{-13}	0.00×10^{0}
300	2.23×10^{-9}	2.82×10^{-9}	0.00×10^{0}	2.00×10^{-11}	1.44×10^{-12}	0.00×10^{0}

Table 6: Results for Example 4.5 with $r = 10^{-12}$.

Let A_0, G_0 and H_0 be 10×10 matrices defined by

$$A_0 = \mathrm{diag}(1, A_{01}, A_{02}, A_{03}, A_{04}, 1),$$
$$G_0 = \mathrm{diag}(10^{-3}, 0, \cdots, 0, 10^{-2}),$$
$$H_0 = \mathrm{diag}(10^{-2}, 0, \cdots, 0, 10^{-3})$$

where

$$A_{01} = \begin{bmatrix} r_1\cos(\pi/3) & r_1\sin(\pi/3) \\ -r_1\sin(\pi/3) & r_1\cos(\pi/3) \end{bmatrix}, \quad A_{02} = \begin{bmatrix} r_2\cos(7\pi/5) & r_2\sin(7\pi/5) \\ -r_2\sin(7\pi/5) & r_2\cos(7\pi/5) \end{bmatrix},$$
$$A_{03} = \begin{bmatrix} r_3\cos(\pi/4) & r_3\sin(\pi/4) \\ -r_3\sin(\pi/4) & r_3\cos(\pi/4) \end{bmatrix}, \quad A_{04} = \begin{bmatrix} r_4\cos(\pi/8) & r_4\sin(\pi/8) \\ -r_4\sin(\pi/8) & r_4\cos(\pi/8) \end{bmatrix},$$

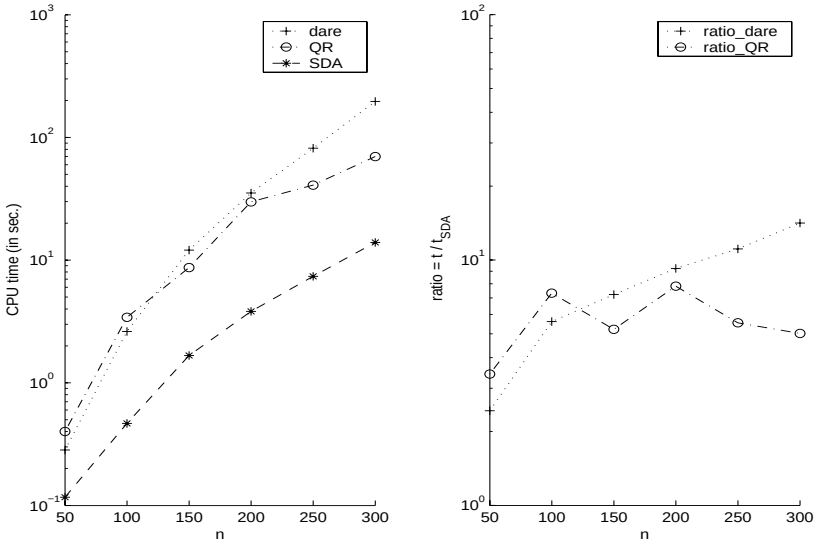


Figure 1: The comparison of CPU times with $r = 1$.

and

$$r_1 = 1 + 3 \times 10^{-15}, \quad r_2 = r_3 = 1 + 10^{-4}, \quad r_4 = 1 + 10^{-5}.$$

Let

$$M_0 = \begin{bmatrix} A_0 & 0 \\ -H_0 & I \end{bmatrix}, \quad L_0 = \begin{bmatrix} I & G_0 \\ 0 & A_0^T \end{bmatrix},$$

$$V_1 = \mathrm{diag}(0, 0.5, \cdots, 0.5, 0), \quad V_2 = \mathrm{diag}(0, 1.5, \cdots, 1.5, 0),$$

and define nonsingular matrices Y_1 and Z_1 by

$$Y_1 = \begin{bmatrix} I & A_0 V_1 (I + H_0 V_1)^{-1} \\ 0 & (I + H_0 V_1)^{-1} \end{bmatrix}, \quad Z_1 = \begin{bmatrix} I & -V_1 \\ 0 & I \end{bmatrix}.$$

A simple calculation gives

$$Y_1 M_0 Z_1 = \begin{bmatrix} A_1 & 0 \\ -H_1 & I \end{bmatrix} \equiv M_1, \quad Y_1 L_0 Z_1 = \begin{bmatrix} I & G_1 \\ 0 & A_1^T \end{bmatrix} \equiv L_1,$$

where $A_1 = A_0(I + V_1 H_0)^{-1}$, $H_1 = (I + H_0 V_1)^{-1} H_0$, and $G_1 = (G_0 - V_1) + A_0 V_1 (I +$

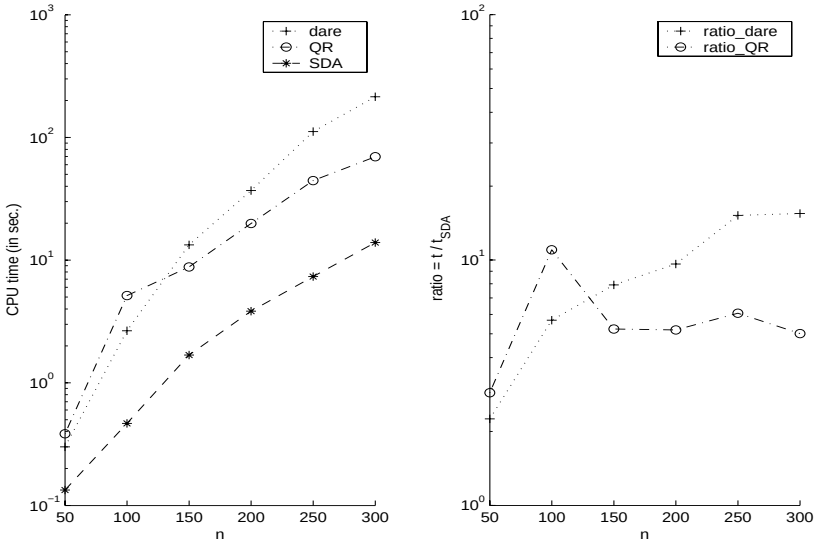


Figure 2: The comparison of CPU times with $r = 10^{-12}$.

$H_0V_1)^{-1}A_0^T$. Furthermore, if we define nonsingular matrices Y_2 and Z_2 by

$$Y_2 = \begin{bmatrix} (I + G_1V_2)^{-1} & 0 \\ -A_1^TV_2(I + G_1V_2)^{-1} & I \end{bmatrix}, \quad Z_2 = \begin{bmatrix} I & 0 \\ V_2 & I \end{bmatrix},$$

then it follows that

$$Y_2M_1Z_2 = \begin{bmatrix} A_2 & 0 \\ -H_2 & I \end{bmatrix} \equiv M_2, \quad Y_2L_1Z_2 = \begin{bmatrix} I & G_2 \\ 0 & A_2^T \end{bmatrix} \equiv L_2,$$

where $A_2 = (I + G_1V_2)^{-1}A_1$, $H_2 = -H_1 + V_2 - A_1^TV_2(I + G_1V_2)^{-1}A_1$, and $G_2 = (I + G_1V_2)^{-1}G_1$. Let $G_2 = B_2B_2^T \geq 0$ and $H_2 = C_2^TC_2 \geq 0$ be the FRD, respectively. Then the system matrices are given by

$$A := U^TA_2U, \quad B := U^TB_2, \quad C := C_2U, \quad R := I.$$

where $U := I - 2uu^T$ with $u = [1, 1, \cdots, 1]^T/\sqrt{10} \in \mathcal{R}^{10}$. It is easy to check that $\min\{||\lambda| - 1| : \lambda \text{ is an eigenvalue of } (M, L)\} \approx 3 \times 10^{-15}$, where $M = \begin{bmatrix} A & 0 \\ -C^TC & I \end{bmatrix}$ and $L = \begin{bmatrix} I & BB^T \\ 0 & A^T \end{bmatrix}$. The numerical results are shown in Table 7.

	dare	QR	SDA
Residual	*	5.68×10^{-5}	6.01×10^{-13}
Iter. no.	-	54	54

Table 7: Results for Example 4.6.

Example 4.7. For $r > 0$, consider a parameterized symplectic pair $(M(r), L(r))$ with

$$A(r) := \begin{bmatrix} 0.4323 & -0.2582 & -1.2863 & 1.8430 & 0.2553 & -0.2746 \\ 0.5969 & -1.8618 & 0.0046 & 0.7127 & 0.3544 & 1.7583 \\ -0.8750 & -1.5715 & -1.3551 & 0.4912 & 0.9922 & 2.1640 \\ -1.0347 & -1.1935 & -0.3797 & 0.8341 & 0.7323 & 1.8743 \\ -0.2771 & -0.8410 & 1.1405 & -1.3839 & -0.2333 & -0.3544 \\ -0.8080 & 0.9526 & 1.2224 & 1.2405 & -1.5662 & 1.5694 \end{bmatrix},$$

$$G(r) := B_2 B_2^T - \frac{1}{r^2} B_1 B_1^T, \quad H(r) := C_1^T C_1,$$

where

$$B_1 = \begin{bmatrix} 0.3447 & 0.6321 & -0.4592 & 1.0773 & 0.2610 & 1.3565 \\ 1.7938 & -0.9404 & -1.1726 & 0.3441 & -0.1703 & -0.1008 \\ 0.6840 & 0.4660 & 1.0479 & 0.1899 & -1.0075 & -0.4529 \\ 0.7424 & 0.6171 & -1.7952 & -0.0011 & 1.7101 & -0.5320 \\ -0.6319 & 0.8059 & -0.6623 & 0.4091 & 0.7990 & 1.4504 \\ -1.7719 & 0.0055 & 0.6855 & 0.0057 & -0.2926 & -0.1119 \end{bmatrix},$$

$$B_2 = \begin{bmatrix} 0.3107 & -0.4471 & 0.1384 & 0.7207 & -1.3962 & -0.7315 \\ 0.5037 & -0.9720 & 0.7164 & -0.3462 & 0.3193 & 1.6300 \\ -1.5449 & -3.0129 & 1.2720 & -1.8523 & -0.4305 & 0.0600 \\ 0.6068 & 0.6410 & 0.1884 & -0.4436 & -1.5227 & -0.1858 \\ 0.2213 & -1.0175 & 0.5326 & 0.2597 & 0.0057 & -0.4042 \\ -0.9153 & 0.1943 & 0.6435 & -1.1077 & -0.1157 & 0.6489 \end{bmatrix},$$

and

$$C_1 = \begin{bmatrix} -2.2752 & 2.1534 & 0.9038 & -1.8451 & 1.4674 & 1.0841 \\ -0.4996 & -1.0463 & 0.6970 & 1.7412 & -1.5000 & -1.6086 \\ 1.7526 & -0.5329 & -1.0929 & -0.6429 & 0.0580 & 1.2661 \\ 0.9504 & 0.4575 & -0.3857 & 1.1104 & 0.1943 & 0.1205 \\ 1.5133 & -0.6674 & 0.5427 & -0.8445 & -1.2548 & 1.3334 \\ -0.7063 & 1.1925 & -0.0400 & 0.4600 & -1.5304 & -0.4101 \end{bmatrix}.$$

This example involves H_∞ norm computation (details in [82]). By applying the bisection method, we can obtain the smallest $r^* \approx 1.08324$ such that $X(r^*) := Ric(M(r^*), L(r^*))$ exists, $I + G(r^*)X(r^*)$ is invertible, and $X(r^*) \geq 0$. Moreover, a pair of real eigenvalues $(\lambda, \frac{1}{\lambda}) = (-0.9999999999998726, -1.000000000000128)$ of the symplectic matrix pair $(M(r^*), L(r^*))$ approach to the unit circle with the distance 1.2745×10^{-13}. The numerical results are given in Table 8. As the s.p.s.d. assumption for $G(r^*)$ is violated for this Example, `dare` cannot be used. This leads to the '$*$' in Table 8.

	dare	QR	SDA
Residual	$*$	5.07×10^{-8}	1.29×10^{-13}
Iter. no.	-	37	22

Table 8: Results for Example 4.7.

Example 4.8. The following example is identical to Example 2.1 of [1], which has been presented originally in [75, Example 2] and [115]. This is an example of stabilizable-detectable, but uncontrollable-unobservable data. We have the following system matrices:

$$A = \begin{bmatrix} 4 & 3 \\ -\frac{9}{2} & -\frac{7}{2} \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad R = \delta, \quad H = \begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix}.$$

The stabilizing solution is

$$X = \frac{1 + \sqrt{1 + 4\delta}}{2} \begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix}.$$

The parameter of R was introduced in [115] to construct ill-conditioned DAREs. Small values for δ will not affect the condition number of DARE much while it grows with increasing values δ. The numerical results with $\delta = 1$ and $\delta = 10^6$ are shown in Table 9.

		`dare`	QR	SDA
$\delta = 1$	Residual	1.34×10^{-14}	8.57×10^{-15}	1.66×10^{-14}
	Rel. err.	9.22×10^{-16}	8.54×10^{-16}	1.46×10^{-16}
	Iter. no.	-	7	6
$\delta = 10^6$	Residual	8.86×10^{-10}	3.62×10^{-5}	5.58×10^{-10}
	Rel. err.	3.98×10^{-11}	1.39×10^{-6}	2.75×10^{-12}
	Iter. no.	-	17	16

Table 9: Results for Example 4.8.

The following three examples come from proportional-plus-integral (PI) control problems. The design includes the original system with coefficient matrices A_1, B_1, C_1, Q_1 and R_1. Additionally, there are r error integrators that are concatenated with the original system. The coefficient matrices of the DARE to be solved is

$$A = \begin{bmatrix} A_1 & 0 \\ -C_1 & I_r \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \quad H = \begin{bmatrix} C_1^T Q_1 C_1 & 0 \\ 0 & Q_2 \end{bmatrix}, \quad R = R_1.$$

Example 4.10. This example is identical to Example 1.11 of [1], which has been presented originally in [104, Section 1.2.2], [54]. The actual data are defined by

$$A_1 = \begin{bmatrix} 0 & 0 \\ I_4 & 0 \\ 0 & A_{22} \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 0.222 & 0.778 & 0 & 0 & 0 \\ 0.4 & 0 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & 1.372 & -0.47 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.098 & 0 \end{bmatrix}^T,$$

and

$$C_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 15 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 7 & -5.357 & -3.943 \end{bmatrix},$$

$$Q_1 = Q_2 = \begin{bmatrix} 0.5 & 0 \\ 0 & 5 \end{bmatrix}, \quad R_1 = \begin{bmatrix} 400 & 0 \\ 0 & 700 \end{bmatrix}.$$

The numerical results are given in Table 10.

	`dare`	QR	SDA
Residual	2.68×10^{-10}	2.38×10^{-4}	1.64×10^{-11}
Iter. no.	-	9	8

Table 10: Results for Example 4.10.

Comments

The Tables show that the approximate solutions from the SDA is either as accurate as or more accurate than those from the QR-SWAP method and `dare`. For the examples considered, the SDA converges to a comparable accuracy, in the same number of iterations or in one less iteration, when compared to the QR-SWAP method. The graphs show the relative efficiencies more clearly, for examples with a parameter reflecting varying degrees of difficulty or conditioning. The efficiency ratios "ratio_dare" and "ratio_QR" stay between 3 and 10. The real ratios should be bigger, as solutions from the SDA are generally more accurate. For example, the residuals and relative errors for SDA for Example 4.5 are virtually zero. Notice that several problems investigated are extremely ill-conditioned. Others have eigenvalues extremely close to the unit circle, numerically violating the assumptions of our theory. The SDA solves them efficiently and accurately without failure.

Example 4.7 comes from an application of the SDA in H_∞ norm computation. The family of examples is dependent on the parameter r, which we would like to minimize

before some stabilizability, detectability and s.p.s.d. constraints are violated. The minimization in this example was carried out by bisection. For r near its minimum, the DARE involved becomes ill-conditioned. This challenging problem was solved in 22 iterations to near machine accuracy. For QR-SWAP, the residual converged to around 10^{-8} in 27 iterations and did not improve further even after 100 more iterations. This behavior, and the similar behavior in Example 4.6, illustrate the importance of the SSF property and the consequent superior convergence, in addition to the better operation count.

There seemed to be a lack of numerical results involving the doubling algorithms, which might have led to the neglect of this class of methods. The preservation of (S) and (D) properties in Lemma 2.1, the convergence results in Theorems 2.3 and 2.4, the superior operation count and the above numerical examples suggest that the neglect has been unjustified.

5 Numerical Experiments for P-DAREs

Similar convention as in Section 4 is utilized in this Section, except the PQZ method [33, 62] replaces `dare` for P-DAREs. In PQZ, QZ decompositions are applied to a $2pn \times 2pn$ matrix pair containing the p matrix pairs defining the P-DAREs. This makes PQZ extremely unattractive in terms of operation counts and efficiency, comparing with the other methods. Also, the method failed when the periodicity p is greater than 10, because of difficulties in the deflation processes involved [33, 62]. Thus, the superiority of the SSCA+SDA for P-DAREs are even more marked than that shown in Section 4 for DAREs. In terms of QR-SWAP for P-DAREs [19, 20], this superiority may be explained by the accumulation of errors in the QR-SWAP method, due to the relative lack of structure preserving properties and the consequent slower convergence.

For the residual of approximate solutions $\{\widetilde{X}_j\}_{j=1}^{p}$, the PQZ method produces p residuals r_j for each $\widetilde{X}_j$ as defined in

$$r_j = \|A_j^T \widetilde{X}_j (I + G_j \widetilde{X}_j)^{-1} A_j + H_j - \widetilde{X}_{j-1}\|_F.$$

The total residual is thus defined as

$$\text{Residual } = \left(\sum_{j=1}^{p} r_j^2 \right)^{1/2}.$$

The situation is different in QR-SWAP and SSCA+SDA, as these methods solve for $\widetilde{X}_p$ via a collapsed matrix pair, generating a residual r_p as in the DARE case. The other $\widetilde{X}$s are obtained through substitutions using the Riccati equation

$$\widetilde{X}_{j-1} = A_j^T \widetilde{X}_j (I + G_j \widetilde{X}_j)^{-1} A_j + H_j.$$

This substitution process generates very little error and virtually no residuals. Consequently, it is difficult to compare the residuals of the PQZ solution with others. From the definition of the residuals and numerical experience, we consider a PQZ solution as equivalent in accuracy if its residual is approximately $\sqrt{p}$ times the residuals from QR-SWAP or the SSCA+SDA. For the Tables in the following examples, data for various methods are lists in columns with obvious headings. The heading "PQZ" is for the periodic QZ algorithm [33, 62], "QR" is for the QR-SWAP method in [20], and for simplicity, the SSCA+SDA algorithm is abbreviated to "SDA".

Example 5.1. As in Example 2 of [62], we consider periodic discrete-time algebraic Riccati equations with $n = p = 3$. The system matrices are

$$A_1 = \begin{bmatrix} -3 & 2 & 9 \\ 0 & 0 & -4 \\ 3 & -2 & 3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 6 & -3 & 0 \\ 4 & -2 & 2 \\ 2 & -1 & 4 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & -3 & -3 \\ 4 & -15 & -3 \\ -2 & 9 & 1 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad R_1 = R_3 = 1, \quad R_2 = 2,$$

$$H_j = e_j e_j^T, \quad j = 1, 2, 3,$$

where e_j denotes the jth column of the identity matrix. The numerical results are reported in Table 11.

	PQZ	QR	SDA
Residual	1.48×10^{-6}	1.66×10^{-7}	2.18×10^{-8}
Iter. no.	-	5	4

Table 11: Results for Example 5.1.

Example 5.2. In [126], the authors considered an optimal periodic output feedback control problem with $n = 4$ and $p = 120$. This periodic discrete-time model was generated from a continuous-time linearized state space model of a spacecraft system [100]. For $j = 1, \ldots, p$, the system matrices are

$$A_j = \begin{bmatrix} 0.9506860 & 0.0429866 & 0.4827320 & -2.5564383 \\ -0.0409684 & 0.9721628 & 1.3617382 & 0.5081454 \\ -0.0122736 & 0.0363280 & -0.8671394 & -0.6014295 \\ -0.0346225 & -0.0072209 & 0.3203622 & -0.8456626 \end{bmatrix},$$

$$B_j = 10^{-5} \begin{bmatrix} 0.2220925 \\ -0.1300536 \\ 0.1877217 \\ -0.0271167 \end{bmatrix} \cos(\omega_0 jT) + 10^{-5} \begin{bmatrix} 0.5035620 \\ 0.4241087 \\ 0.1218290 \\ 0.3583826 \end{bmatrix} \sin(\omega_0 jT),$$

$$C_j = \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad R_j = 10^{-11}.$$

where $\omega_0 = 0.00103448$ rad/s is the orbital frequency and $T = 2\pi/(\omega_0 p)$ is the sampling period. The numerical results are reported in Table 12.

Example 5.3. In this example, we tested the three methods on the randomly generated periodic matrix pairs $\{(M_j, L_j)\}_{j=1}^{p}$. Entries of A_j are distributed normally in the interval $[-2, 2]$, and entries of matrices B_j , C_j are distributed normally in the interval $[-1, 1]$ $(j = 1, \ldots, p)$. We set $\text{rank}(B_j) = \text{rank}(C_j) = 0.7n$, for all j.

Figure 3 reports the comparison of CPU times for $n = 50, 100, 150, 200, 250, 300$, all with $p = 8$. Figure 4 reports the comparison of CPU times for $p = 4, 8, 16, 32, 64, 128$,

	PQZ	QR	SDA
Residual	*	2.43×10^{-13}	2.00×10^{-14}
CPU time	*	0.133	0.083
Iter. no.	-	3	2

Table 12: Results for Example 5.2 with $n = 4$ and $p = 120$.

with $n = 30$. In these two cases, the numerical results of residuals are shown in Table 13.

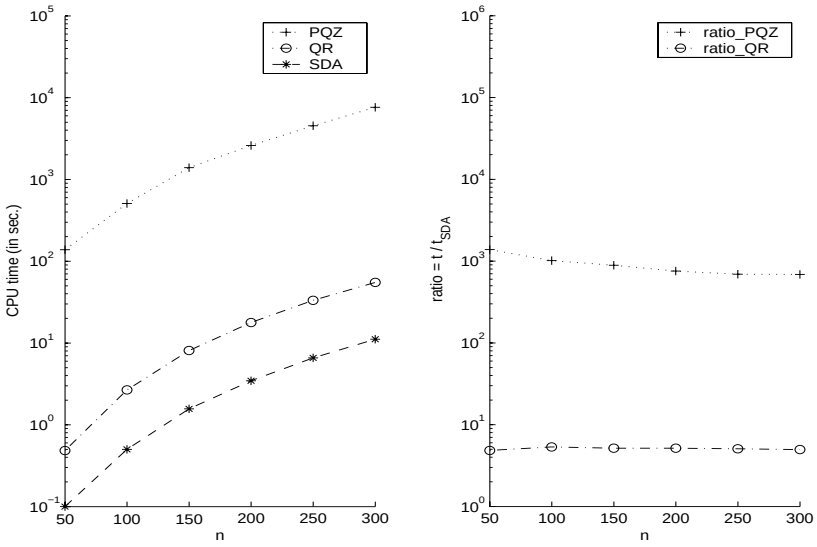


Figure 3: The comparison of CPU times for $n = 50, 100, 150, 200, 250, 300$ and $p = 8$.

Comments

For Example 5.1, the SDA produced the most accurate solution in 3 iterations, as compared to 4 for QR-SWAP. The solution from PQZ is two order of magnitude worse than those from the SDA.

The PQZ method failed for Example 5.2, probably because of the small elements in R_j

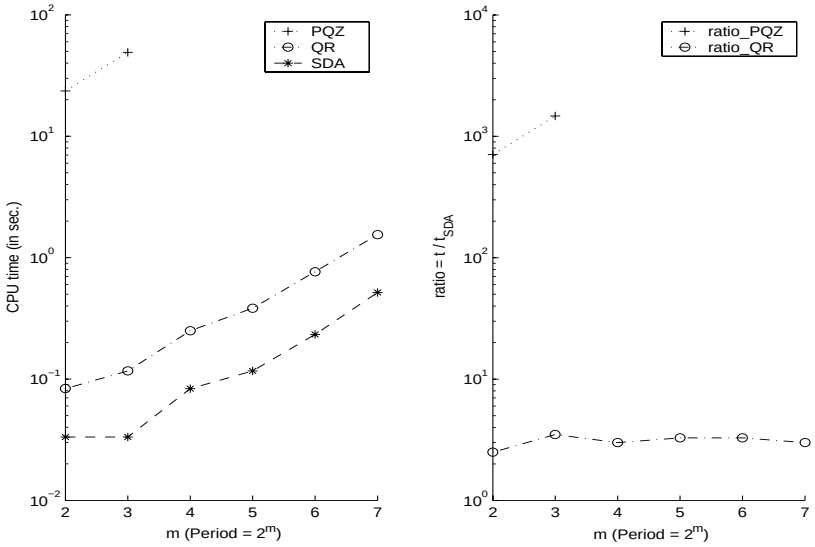


Figure 4: The comparison of CPU times for $p = 2, 4, 8, 16, 32, 128$ and $n = 30$.

and C_j. Near-machine accuracy was achieved by the SDA in 2 iterations. The QR-SWAP method produced a solution one order of magnitude worse in one more iteration.

Example 5.3 contains several randomly generated examples, with varying values of n and p. The graphs show that the SDA is around 3 to 5 times more efficient than QR-SWAP, and around 1000 times more efficient than PQZ, which failed for large values of p. This can be explained by the well-known fact that the shift and deflate approach in PQZ fails when p is large. Notice that the efficiency advantage should be higher because of the smaller residuals for solutions from the SDA.

The examples illustrated that the SDA is more efficient than the PQZ and QR-SWAP algorithms.

6 Conclusions

We conclude this chapter with a summary of results and a few comments.

SDA and QR-SWAP In this chapter, we investigate structure-preserving algorithms (SDA and SSCA) for solving DAREs and P-DAREs and prove the quadratical con-

p	n	res_PQZ	res_QR	res_SDA
8	50	1.29×10^{-5}	3.94×10^{-6}	3.64×10^{-6}
	100	5.72×10^{-4}	4.13×10^{-4}	2.08×10^{-4}
	150	6.03×10^{-3}	4.30×10^{-3}	2.56×10^{-3}
	200	3.47×10^{-2}	2.48×10^{-2}	1.41×10^{-2}
	250	1.21×10^{-1}	9.91×10^{-2}	4.39×10^{-2}
	300	3.42×10^{-1}	2.17×10^{-1}	1.30×10^{-1}
4	30	4.99×10^{-7}	7.42×10^{-7}	2.20×10^{-7}
8		8.56×10^{-7}	9.34×10^{-7}	4.04×10^{-7}
16		$*$	4.13×10^{-7}	2.42×10^{-7}
32		$*$	3.19×10^{-7}	2.03×10^{-7}
64		$*$	1.12×10^{-6}	3.17×10^{-7}
128		$*$	6.51×10^{-7}	2.79×10^{-7}

Table 13: Results of Example 5.3 for various p and n.

vergence of the SDA under assumptions which are weaker than stabilizability and detectability. P-DAREs are first reduced to a DARE by the SSCA. The resulting DARE is then solved by the iterative SDA. The algorithm looks, on the surface, very similar to the QR-SWAP algorithm in [20]. The algorithms SDA and QR-SWAP are obviously closely related, sharing similar theoretical background and convergence analysis. However, there are some important differences in the details. The main difference is in the stronger SSF properties in the SDA, preserving the symplecticity in standard symplectic form as well as the stabilizability and detectability properties, through the iterative process. In addition, the SDA allows the iteration to be carried out with far fewer flops. It is interesting that the swap and collapse steps in QR-SWAP are forward stable numerically, as compared to the structure-preserving steps (SSF) in the SDA are numerically efficient and reliable (recall the inversion of the well-behaved matrix operation $(I + GH)$ with $G,\ H \geq 0$). It may

be the case that smaller errors (in QR-SWAP) can do more harm than larger but structured errors (in SDA). Also, the least squares step at the end of QR-SWAP is not required in the SDA. Together with the difference in operation counts, the SDA seems to be a superior algorithm. Notice that the PQZ algorithm (or its equivalent `dare` for DAREs) is never a competitor to QR-SWAP or the SDA, due to its inferior operation count. Notice also that the SDA performs a lot better for ill-conditioned P-DAREs. In summary, the numerical evidence we have gathered so far indicates that the SDA is an accurate, robust and efficient algorithm for P-DAREs. The algorithm appears to be a sound basis on which a general-purpose algorithm for P-DAREs can be built.

Deficiency in SDA There is one advantage of QR-SWAP over the SDA that we are aware of. Let the state equation $x_{j+1} = A_j x_j + B_j u_j$ be replaced by a descriptor system $E_j x_{j+1} = A_j x_j + B_j u_j$ with a nonsingular but ill-conditioned E_j. The QR-SWAP algorithm should still work while the SDA will founder, because the inversion of E_j is required in the SSF structure.

Parallelism It is easy to see that all the possibilities of parallelism in QR-SWAP [20] exist in SSCA and the SDA. Recall that the swap and collapse procedure in the SDA and QR-SWAP can be carried out simultaneous at different points. For example for P-DAREs with period $p = 4$, we can swap and collapse the first two matrix pairs in parallel to the same operations on the last two, and then swap and collapse the two resulting matrix pairs into the final pair.

Appendix

Proof of Lemma 2.3. Let

$$v^T \widehat{A} = \lambda v^T, \quad |\lambda| \geq 1 \tag{A.1}$$

and

$$v^T \widehat{G} = v^T (G + AG(I + HG)^{-1} A^T) = 0. \tag{A.2}$$

We need to show that $v = 0$ for the stabilizability of $(\widehat{A}, \widehat{B})$, where $\widehat{G} = \widehat{B}\widehat{B}^T \geq 0$ is a FRD. From (A.2), the FRD of $G = BR^{-1}B^T \geq 0$ and the fact that $(G(I + HG))^{-1}$ is s.p.s.d. (two application of the SMWF) follows that

$$v^T B = 0, \quad v^T AB = 0. \tag{A.3}$$

Substituting (2.9) into (A.1), and using the SMWF and (A.3), we obtain

$$v^T \widehat{A} = v^T A(I + GH)^{-1}A = v^T A \left[I - G(I + HG)^{-1}H\right] A = v^T A^2 = \lambda v^T. \tag{A.4}$$

From (A.4) follows that v is a linear combination of two left eigenvectors $\{u_1, u_2\}$ of A corresponding to the eigenvalues $\{\omega_1, \omega_2\}$ $(\omega_1 \neq \omega_2)$, i.e. $u_j^T A = \omega_j u_j^T$, with $\omega_j^2 = \lambda$, $j = 1, 2$. Write

$$v = \alpha_1 u_1 + \alpha_2 u_2. \tag{A.5}$$

Substituting (A.5) into (A.3) and eliminating $\alpha_2\omega_2 u_2 B$, we obtain that $u_1^T B = 0$. From the stabilizability of (A, B) and the relation $u_1^T A = \omega_1 A$, $|\omega_1| \geq 1$, it follows that $u_1 = 0$. Similarly, we can also show that $u_2 = 0$. These imply $v = 0$.

The detectability result of $(\widehat{A}, \widehat{C})$ can be proved similarly. □

Proof of Lemma 3.2. Let

$$v^T \widehat{A}_p = \lambda v^T \ , \quad |\lambda| \geq 1 \tag{A.6}$$

and

$$v^T \widehat{G}_p = v^T [G_p + A_p \widehat{G}_{p-1}(I_{n_{p-1}} + H_p \widehat{G}_{p-1})^{-1} A_p^T] = 0. \tag{A.7}$$

We need to show that $v = 0$ for the stabilizability of $(\widehat{A}_p,\ \widehat{B}_p)$, where $\widehat{G}_p = \widehat{B}_p \widehat{B}_p^T \geq 0$ is a FRD.

It can be shown from (3.9)–(3.11) that the FRDs of $\widehat{G}_{p-1} = \widehat{B}_{p-1}\widehat{B}_{p-1}^T \geq 0$ and $\widehat{G}_{p-1}(I_{n_{p-1}} + H_p \widehat{G}_{p-1})^{-1}$ is s.p.s.d. (two applications of the SMWF). From (A.7) and the FRD of $G_p = B_p R_p^{-1} B_p^T$, it then follows that

$$v^T B_p = 0 \ , \quad v^T A_p \widehat{B}_{p-1} = 0 \ . \tag{A.8}$$

Substituting (3.9)–(3.11) into (A.6) and using the SMWF, we obtain

$$v^T \widehat{A}_p = v^T A_p (I_{n_{p-1}} + \widehat{G}_{p-1} H_p)^{-1} \widehat{A}_{p-1} = v^T A_p \left[I_{n_{p-1}} - \widehat{G}_{p-1} (I_{n_{p-1}} + H_p \widehat{G}_{p-1})^{-1} H_p \right] \widehat{A}_{p-1}$$
$$= v^T A_p \widehat{A}_{p-1} = \lambda v^T. \tag{A.9}$$

Repeating the argument in (A.8)–(A.9) using (3.9)–(3.11), we arrive at

$$v^T A_p \cdots A_1 = \lambda v^T \quad \text{with} \quad |\lambda| \geq 1$$

and

$$v^T B_p = v^T A_p B_{p-1} = \cdots = v^T A_p A_{p-1} \cdots A_2 B_1 = 0.$$

These imply $v = 0$ because of the (P-S) property of the original periodic system. The (P-D) can be proved similarly. □

Chapter 2

Structure-Preserving Doubling Algorithm for CAREs

1 Introduction

In this chapter we investigate a structure-preserving doubling algorithm for the computation of the symmetric positive semi-definite (s.p.s.d.) solution X (i.e. $X \geq 0$) to the continuous-time algebraic Riccati equation (CARE):

$$-XGX + A^T X + XA + H = 0, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$, $X \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{m\times m}$ is symmetric positive definite (or s.p.d.; i.e. $R > 0$), $G = BR^{-1}B^T \geq 0$ and $H = C^T C \geq 0$ with $B \in \mathbb{R}^{n\times m}$ and $C^T \in \mathbb{R}^{n\times p}$ being of full column rank.

Equation (1.1) arises frequently in solving the continuous-time linear optimal control problem:

$$\min_u J = \frac{1}{2}\int_0^\infty (x^T C^T C x + u^T R u)\, dt \qquad \text{subject to} \quad \dot{x} = Ax + Bu. \tag{1.2}$$

The optimal feedback control u^* for (2) is given by

$$u^* = -R^{-1}B^T X x, \tag{1.3}$$

where X is the s.p.s.d. solution to the CARE (1.1). We assume that the pair (A, B) is stabilizable (S) (i.e. if $w^T B = 0$ and $w^T A = \lambda w^T$ for some $\lambda \in \mathbb{C}$, then $\mathrm{Re}(\lambda) < 0$ or $w = 0$) and that the pair (A, C) is detectable (D) (i.e. (A^T, C^T) is stabilizable). Under assumptions (S) and (D), the CARE (1.1) has been proved to possess a unique s.p.s.d. solution [75].

Consider the $2n \times 2n$ Hamiltonian matrix $\mathcal{H}$ associated with the CARE (1.1):

$$\mathcal{H} = \begin{bmatrix} A & -G \\ -H & -A^T \end{bmatrix} \tag{1.4}$$

which satisfies

$$\mathcal{H}J = -J\mathcal{H}^T, \quad J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

with I_n denoting the identity matrix of order n. By (1.4), the CARE (1.1) can be written as

$$\mathcal{H} \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} \Phi, \tag{1.5}$$

where $\Phi \in \mathbb{R}^{n \times n}$ and the spectrum $\sigma(\Phi)$ is on the stable left half plane $\mathbb{C}_-$. Under assumptions (S) and (D), the Hamiltonian matrix $\mathcal{H}$ has exactly n eigenvalues on $\mathbb{C}_-$. If the columns of $[X_1^T, X_2^T]^T$ span the stable invariant subspace of $\mathcal{H}$, then X_1 is nonsingular and $X = X_2 X_1^{-1} \geq 0$ solves the CARE (1.1) (see, e.g., [75, 95]).

A numerically backward stable algorithm `care`, proposed by Laub [75], computes X by applying the QR algorithm with reordering [5, 34, 108] to the eigenvalue problem $\mathcal{H}x = \lambda x$. Unfortunately, the QR algorithm preserves neither the Hamiltonian structure of $\mathcal{H}$ nor the associated splitting of eigenvalues. A structure-preserving algorithm has been proposed by Ammar and Mehrmann [2] which utilizes orthogonal symplectic transformations in computing a basis for the stable invariant subspace of $\mathcal{H}$. A stable symplectic orthogonal method has been suggested by Byers [39] but applied only to systems with single input or output. Many iterative methods have been suggested for solving CAREs over the past 20 years. Newton's method has been applied in extensive literature [51, 61, 71, 91, 103]. A defect correction method for modifying an approximate solution has also been proposed by Mehrmann and Tan [93]. These methods require a good starting approximate solution, and can therefore be regarded as iterative refinement methods, to be combined with other direct methods (see Bunse-Gerstner *et al* [36, 38] or Mehrmann [91] for details). The structure-preserving matrix sign function methods

(MSGM) [9, 13, 14, 15, 16, 40, 41, 50, 64, 102] have been extended by Barraud [10, 11] and Cardiner and Laub [56].

A class of methods, referred to as the doubling algorithms (DA), has attracted much interests in the 70s and 80s (see [3] and the references therein). These methods originate from the fixed-point iteration derived from the discrete-time algebraic Riccati equation (DARE):

$$X_{k+1} = \widehat{A}^T X_k (I + \widehat{G} X_k)^{-1} \widehat{A} + \widehat{H} \ .$$

Instead of producing the sequence $\{X_k\}$, doubling algorithms produce $\{X_{2^k}\}$. CAREs can be tackled after being transformed to DAREs via the Cayley transform. However, the convergence of the algorithm was proven only when $\widehat{A}$ is nonsingular [3], and for $(\widehat{A}, \widehat{G}, \widehat{H})$ which is stabilizable and detectable [69]. DAs were largely forgotten in the past decade. Recently, DAs have been revived for (periodic) DAREs, because of a better theoretical understanding. Stronger convergence results have been proved for $(\widehat{A}, \widehat{G}, \widehat{H})$ under weaker assumptions than stabilizability and detectability (see Section 2 of Chapter 1). Superior numerical results, in comparison to state-of-the-art methods on a wide range of test problems, have been obtained because of the stronger structure-preserving properties and the superior operations count.

In this chapter, we propose a doubling algorithm for CAREs. The CAREs are transformed to DAREs, with the corresponding Hamiltonian matrix transformed into a symplectic matrix pair by the Cayley transform. Nice convergence properties are inherited from the structure-preserving doubling algorithm (SDA) applied to the corresponding DARE. The SDA preserves matrix pairs in SSF which is a stronger property than symplecticity. In the CARE setting, the matrix sign function methods preserve the Hamiltonian structure in $\mathcal{H}$ while the SDA preserves, in each iterative step, the associated symplectic matrix pair $(\widehat{\mathcal{N}}, \widehat{\mathcal{L}})$ in SSF. Although under the influence of numerical errors, the matrix pairs through the SDA retain their stabilizability, detectability as well as eigenstructures (with exactly half of the spectrum being stable; see details in Section 2 of Chapter 1). This stronger structure-preserving property is its main strength and the reason of its accuracy. In Section 4, a modified version of the SDA (SDA_m) is developed, for "doubly

symmetric" DAREs, where $\hat{A}$, $\hat{G} = \hat{H}$ are symmetric and persymmetric. The SDA_m preserves the symplectic and doubly symmetric structures of the DARE, resulting in better accuracy than the SDA. We have extensively tested the SDA against the MSGM and `care`. Numerical results showed that the doubling algorithm for CAREs is competitive and promising.

Finally, it is important to stress that matrix sign functions can be applied to more general Hamiltonian matrices in other applications, such as those from H_∞ control with G and H being indefinite. A scaling strategy [41] may also accelerate its convergence. Also, the SDA requires the transformation of the CARE by the Cayley transform, which requires the estimation of the parameter γ (see §3 below).

2 SDA and Matrix Sign Function Method

In this section we propose a structure-preserving doubling algorithm (SDA) for solving the CARE (1.1) based on the doubling algorithm proposed in Section 2 of Chapter 1. In addition, the well-known structure-preserving matrix sign function methods [9, 13, 14, 15, 16, 40, 41, 50, 64, 102] are also reviewed from the point of view of preserving Hamiltonian structure.

Let $\mathbf{H}$ be the set of $2n \times 2n$ Hamiltonian matrices, i.e.,

$$\mathbf{H} = \left\{ \mathcal{H} \left| \mathcal{H} = \begin{bmatrix} A & -G \\ -H & -A^T \end{bmatrix} ;\ A, H, G \in \mathbb{R}^{n\times n};\ H, G \geq 0 \right. \right\}. \tag{2.1}$$

Note that if $\mathcal{H} \in \mathbf{H}$ then $\mathcal{H}J = -J\mathcal{H}^T$. We call a $2n \times 2n$ matrix pair $(\mathcal{N}, \mathcal{L})$ symplectic if $\mathcal{N}J\mathcal{N}^T = \mathcal{L}J\mathcal{L}^T$. Let $\mathbf{S}$ be the set of $2n \times 2n$ symplectic matrix pairs in the standard symplectic form (SSF):

$$\mathbf{S} = \left\{ (\widehat{\mathcal{N}}, \widehat{\mathcal{L}}) \left| \widehat{\mathcal{L}} = \begin{bmatrix} I & \widehat{G} \\ 0 & \widehat{A}^T \end{bmatrix},\ \widehat{\mathcal{N}} = \begin{bmatrix} \widehat{A} & 0 \\ -\widehat{H} & I \end{bmatrix} ;\ \widehat{A}, \widehat{H}, \widehat{G} \in \mathbb{R}^{n\times n};\ \widehat{G}, \widehat{H} \geq 0 \right. \right\}. \tag{2.2}$$

It is easily seen that symplecticity is weaker than symplecticity in SSF. Our proposed algorithm preserves the stronger structure and gives rise to better numerical performance.

We shall show how the CARE (1.1), associated with the corresponding Hamiltonian matrix

$$\mathcal{H} \equiv \begin{bmatrix} A & -G \\ -H & -A^T \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^T \\ -C^TC & -A^T \end{bmatrix} \in \mathbb{R}^{2n\times 2n},$$

can be transformed to an equivalent DARE.

By using the Cayley transform with some appropriate $\gamma > 0$, the Hamiltonian matrix $\mathcal{H}$ can be transformed to a symplectic matrix pair $(\mathcal{N}, \mathcal{L}) \equiv (\mathcal{H} + \gamma I, \mathcal{H} - \gamma I)$ [91, 92]. In the following, we construct an equivalence transformation from $(\mathcal{N}, \mathcal{L})$ to a symplectic matrix pair $(\widehat{\mathcal{N}}, \widehat{\mathcal{L}}) \in \mathbf{S}$.

Let

$$A_\gamma \equiv A - \gamma I, \quad \bar{A}_\gamma \equiv A + \gamma I.$$

Starting from

$$\mathcal{N} = \begin{bmatrix} \bar{A}_\gamma & -G \\ -H & -A_\gamma^T \end{bmatrix}, \quad \mathcal{L} = \begin{bmatrix} A_\gamma & -G \\ -H & -\bar{A}_\gamma^T \end{bmatrix},$$

we choose a $\gamma > 0$ such that the matrices A_γ and $A_\gamma + GA_\gamma^{-T}H$ are well-conditioned (see Section 3 later for details). To transform the symplectic matrix pair $(\mathcal{N}, \mathcal{L})$ to $(\widehat{\mathcal{N}}, \widehat{\mathcal{L}}) \in \mathbf{S}$, let

$$T_1 \equiv \begin{bmatrix} A_\gamma^{-1} & 0 \\ HA_\gamma^{-1} & I \end{bmatrix}, \quad T_2 \equiv \begin{bmatrix} I & 0 \\ 0 & (-HA_\gamma^{-1}G - A_\gamma^T)^{-1} \end{bmatrix}, \quad T_3 \equiv \begin{bmatrix} I & A_\gamma^{-1}G \\ 0 & I \end{bmatrix}.$$

Simple calculations produce

$$\widehat{\mathcal{N}} = \begin{bmatrix} \widehat{A} & 0 \\ -\widehat{H} & I \end{bmatrix} = T_3T_2T_1\mathcal{N} = T_3T_2 \begin{bmatrix} A_\gamma^{-1}\bar{A}_\gamma & -A_\gamma^{-1}G \\ HA_\gamma^{-1}\bar{A}_\gamma - H & -HA_\gamma^{-1}G - A_\gamma^T \end{bmatrix}$$

$$= T_3 \begin{bmatrix} A_\gamma^{-1}\bar{A}_\gamma & -A_\gamma^{-1}G \\ (-HA_\gamma^{-1}G - A_\gamma^T)^{-1}(HA_\gamma^{-1}\bar{A}_\gamma - H) & I \end{bmatrix}$$

and

$$\widehat{\mathcal{L}} = \begin{bmatrix} I & \widehat{G} \\ 0 & \widehat{A}^T \end{bmatrix} = T_3T_2T_1\mathcal{L} = T_3T_2 \begin{bmatrix} I & -A_\gamma^{-1}G \\ 0 & -HA_\gamma^{-1}G - \bar{A}_\gamma^T \end{bmatrix}$$

$$= T_3 \begin{bmatrix} I & -A_\gamma^{-1}G \\ 0 & (-HA_\gamma^{-1}G - A_\gamma^T)^{-1}(-HA_\gamma^{-1}G - \bar{A}_\gamma^T) \end{bmatrix},$$

where

$$\begin{aligned}
\widehat{A} &= (\bar{A}_\gamma + GA_\gamma^{-T}H)(A_\gamma + GA_\gamma^{-T}H)^{-1}, \\
\widehat{G} &= -A_\gamma^{-1}G + A_\gamma^{-1}G(A_\gamma^T + HA_\gamma^{-1}G)^{-1}(\bar{A}_\gamma^T + HA_\gamma^{-1}G), \\
\widehat{H} &= (A_\gamma^T + HA_\gamma^{-1}G)^{-1}(HA_\gamma^{-1}\bar{A}_\gamma - H).
\end{aligned}$$

Note that $\mathcal{L}^{-1}\mathcal{N} = \widehat{\mathcal{L}}^{-1}\widehat{\mathcal{N}}$. Since $\bar{A}_\gamma = A_\gamma + 2\gamma I$, it follows that

$$\widehat{A} = I + 2\gamma(A_\gamma + GA_\gamma^{-T}H)^{-1}, \tag{2.3}$$

$$\widehat{G} = 2\gamma A_\gamma^{-1}G(A_\gamma^T + HA_\gamma^{-1}G)^{-1}, \tag{2.4}$$

$$\widehat{H} = 2\gamma(A_\gamma^T + HA_\gamma^{-1}G)^{-1}HA_\gamma^{-1}. \tag{2.5}$$

Then we obtain the desired symplectic matrix pair in SSF, i.e.,

$$(\widehat{\mathcal{N}}, \widehat{\mathcal{L}}) \equiv \left(\begin{bmatrix} \widehat{A} & 0 \\ -\widehat{H} & I \end{bmatrix}, \begin{bmatrix} I & \widehat{G} \\ 0 & \widehat{A}^T \end{bmatrix} \right) \in \mathbf{S},$$

where $\widehat{A}$, $\widehat{G}$ and $\widehat{H}$ are given by (2.3)–(2.5). The DARE associated with the symplectic matrix pair $(\widehat{\mathcal{N}}, \widehat{\mathcal{L}})$ in SSF is

$$X = \widehat{A}^T X(I + \widehat{G}X)^{-1}\widehat{A} + \widehat{H} \tag{2.6}$$

on which the efficient SDA, proposed in Section 2 of Chapter 1, can be applied. Note that X is the unique s.p.s.d. solution to the above DARE as well as the CARE (1.1). Moreover, in Theorems 1 and 2 of [70], the pairs $(\widehat{A}, \widehat{B})$ and $(\widehat{A}, \widehat{C})$ are proven to be stabilizable and detectable, respectively, where the matrices $\widehat{G} = \widehat{B}\widehat{B}^T$ and $\widehat{H} = \widehat{C}^T\widehat{C}$ are full rank decompositions (FRD).

Using (2.3)–(2.5) to transform the CARE (1.1) to an equivalent DARE (2.6) with the associated symplectic matrix pair $(\widehat{\mathcal{N}}, \widehat{\mathcal{L}})$ in SSF, the SDA proposed in Section 2 of Chapter 1 can then be modified to the following algorithm for CAREs: (with Im denoting the imaginary axis)

Structure-Preserving Doubling Algorithm (SDA):

Input: $\mathcal{H} = \begin{bmatrix} A & -G \\ -H & -A^T \end{bmatrix} \in \mathbf{H}$ with $\sigma(\mathcal{H}) \cap \mathrm{Im} = \emptyset$; ϵ

Output: the stabilizing solution $X = X^T \geq 0$ to the CARE (1.1).

Find an appropriate value $\hat{\gamma} > 0$.

Compute $\widehat{A}_0 \leftarrow I + 2\hat{\gamma}(A_{\hat{\gamma}} + GA_{\hat{\gamma}}^{-T}H)^{-1}$, $\widehat{G}_0 \leftarrow 2\hat{\gamma}A_{\hat{\gamma}}^{-1}G(A_{\hat{\gamma}}^T + HA_{\hat{\gamma}}^{-1}G)^{-1}$,

$\widehat{H}_0 \leftarrow 2\hat{\gamma}(A_{\hat{\gamma}}^T + HA_{\hat{\gamma}}^{-1}G)^{-1}HA_{\hat{\gamma}}^{-1}$, $j \leftarrow 0$;

Do until convergence:

Compute $\widehat{A}_{j+1} \leftarrow \widehat{A}_j(I + \widehat{G}_j\widehat{H}_j)^{-1}\widehat{A}_j$, $\widehat{G}_{j+1} \leftarrow \widehat{G}_j + \widehat{A}_j\widehat{G}_j(I + \widehat{H}_j\widehat{G}_j)^{-1}\widehat{A}_j^T$,

$\widehat{H}_{j+1} \leftarrow \widehat{H}_j + \widehat{A}_j^T(I + \widehat{H}_j\widehat{G}_j)^{-1}\widehat{H}_j\widehat{A}_j$, $j \leftarrow j + 1$;

If $\|\widehat{H}_j - \widehat{H}_{j-1}\| \leq \epsilon\|\widehat{H}_j\|$, Stop;

End

Set $X \leftarrow \widehat{H}_j$.

Convergence of SDA

Let $\widehat{\mathcal{N}} = \begin{bmatrix} \widehat{A} & 0 \\ -\widehat{H} & I \end{bmatrix}$, $\widehat{\mathcal{L}} = \begin{bmatrix} I & \widehat{G} \\ 0 & \widehat{A}^T \end{bmatrix}$, where $\widehat{G} = \widehat{G}^T$, $\widehat{H} = \widehat{H}^T$. Suppose $\widehat{\mathcal{N}} - \lambda\widehat{\mathcal{L}}$ has no eigenvalues on the unit circle and there exist nonsingular Q, Z such that

$$Q\widehat{\mathcal{N}}Z = \begin{bmatrix} J_s & 0 \\ 0 & I \end{bmatrix}, \quad Q\widehat{\mathcal{L}}Z = \begin{bmatrix} I & 0 \\ 0 & J_s \end{bmatrix} \tag{2.7}$$

where the spectrum $\lambda(J_s) \in O_s \equiv \{\lambda : |\lambda| < 1\}$. In the following we quote the convergence results for the SDA algorithm from Section 2 of Chapter 1.

Theorem 2.1. *Let* $\widehat{\mathcal{N}} = \begin{bmatrix} \widehat{A} & 0 \\ -\widehat{H} & I \end{bmatrix}$ *and* $\widehat{\mathcal{L}} = \begin{bmatrix} I & \widehat{G} \\ 0 & \widehat{A}^T \end{bmatrix}$, *where* $\widehat{G} = \widehat{G}^T$, $\widehat{H} = \widehat{H}^T$. *Suppose* $\widehat{\mathcal{N}} - \lambda\widehat{\mathcal{L}}$ *has no eigenvalues on the unit circle and there exist nonsingular* Q, Z *such that (2.7) holds. Denote* $Z = \begin{bmatrix} Z_1 & Z_3 \\ Z_2 & Z_4 \end{bmatrix}$, $Z_i \in \mathbb{R}^{n\times n}$ *for* $i = 1, 2, 3, 4$. *If* Z_1 *and* Z_4 *are invertible, then the sequences* $\{\widehat{A}_j,\ \widehat{H}_j,\ \widehat{G}_j\}$ *computed by the SDA algorithm satisfy*

(i) $\|\widehat{A}_j\| = O(\|J_s^{2^j}\|) \to 0$ *as* $j \to \infty$,

(ii) $\widehat{H}_j \to X$, *where* X *solves the DARE (2.6):*

$$X = \widehat{A}^T X(I + \widehat{G}X)^{-1}\widehat{A} + \widehat{H},$$

(iii) $\widehat{G}_j \to Y$, *where* Y *solves the dual DARE*

$$Y = \widehat{A}Y(I + \widehat{H}Y)^{-1}\widehat{A}^T + \widehat{G}. \tag{2.8}$$

Moreover, the convergence rate in (i)–(iii) above is $O\left(|\lambda_n|^{2^j}\right)$, *where* $|\lambda_1| \le \cdots \le |\lambda_n| < 1 < |\lambda_n|^{-1} \le \cdots \le |\lambda_1|^{-1}$ *with* λ_i, λ_i^{-1} *being the eigenvalues of* $\widehat{\mathcal{N}} - \lambda\widehat{\mathcal{L}}$ *(including* 0 *and* ∞*).*

NG HUA UNIVERSITY

The following Lemma proves that the stabilizability and detectability properties are preserved by the SDA throughout its iterative process.

Lemma 2.2. *The stabilizability of* $(\widehat{A}, \widehat{B})$ *implies that* $(\widehat{A}_j, \widehat{B}_j)$ *is stabilizable, where* $\widehat{G}_j = \widehat{B}_j\widehat{B}_j^T \ge 0$ *is a FRD of* $\widehat{G}_j$ *for each* $j \ge 1$. *The detectability of* $(\widehat{A}, \widehat{C})$ *implies that* $(\widehat{A}_j, \widehat{C}_j)$ *is detectable, where* $\widehat{H}_j = \widehat{C}_j^T\widehat{C}_j \ge 0$ *is a FRD of* $\widehat{H}_j$ *for each* $j \ge 1$.

Theorem 2.3. *Let* $\widehat{\mathcal{N}} = \begin{bmatrix} \widehat{A} & 0 \\ -\widehat{H} & I \end{bmatrix}$ *and* $\widehat{\mathcal{L}} = \begin{bmatrix} I & \widehat{G} \\ 0 & \widehat{A}^T \end{bmatrix}$, *where the matrices* $\widehat{G} = \widehat{B}\widehat{B}^T \ge 0$ *(FRD) and* $\widehat{H} = \widehat{C}^T\widehat{C} \ge 0$ *(FRD). Assume that* $(\widehat{A}, \widehat{B})$ *is stabilizable and* $(\widehat{A}, \widehat{C})$ *is detectable. Then the sequences* $\{\widehat{A}_j, \widehat{H}_j, \widehat{G}_j\}$ *computed by the SDA satisfy (i), (ii), (iii) as in Theorem 2.1.*

Remark. Theorem 2.1 directly proves, under the assumptions that $\widehat{\mathcal{N}} - \lambda\widehat{\mathcal{L}}$ have no unit modulo eigenvalues and Z_1, Z_4 are invertible, that the sequences $\{\widehat{A}_j, \widehat{H}_j, \widehat{G}_j\}$ generated by the SDA converge to zero and the unique s.p.s.d. solutions of the DAREs in (2.6) and (2.8), respectively. Lemma 2.2 shows the preservation of stabilizability and detectability of the iterates $(\widehat{A}_j, \widehat{G}_j, \widehat{H}_j)$ generated by the SDA. Furthermore, in Theorem 2.3, we see that the assumptions in Theorem 2.1 are weaker than conditions (S) and (D). This distinction of preserving the symplectic structure in SSF, as well as the difference in operation counts, are responsible for the superior performance of the SDA.

On the other hand, for a given $\mathcal{H} = \begin{bmatrix} A & -G \\ -H & -A^T \end{bmatrix} \in \mathbf{H}$ with $\sigma(\mathcal{H}) \cap \mathrm{Im} = \emptyset$, the matrix sign function of $\mathcal{H}$ can also be used to develop a structure-preserving method for computing the stabilizing solution of CARE (1.1). A thorough discussion and the details of practical implementation are given in [41, 91]. The main MSGM algorithm is described as follows. Other modified versions can be found in [6, 10, 11, 42, 56] and references therein.

Matrix sign function algorithm: [9, 13, 14, 15, 16, 40, 41, 50, 64, 102]

`Input`: $\mathcal{H} = \begin{bmatrix} A & -G \\ -H & -A^T \end{bmatrix} \in \mathbf{H}$ with $\sigma(\mathcal{H}) \cap \mathrm{Im} = \emptyset$; ϵ.

`Output`: the stabilizing solution $X = X^T \geq 0$ to the CARE (1.1).

`Let` $\mathcal{H}_0 \leftarrow \mathcal{H}$, $j \leftarrow 0$.

`Do` until convergence:

`Compute` $\mathcal{H}_{j+1} \leftarrow \frac{1}{2}(\mathcal{H}_j + \mathcal{H}_j^{-1})$, $j \leftarrow j + 1$;

If $\|\mathcal{H}_j - \mathcal{H}_{j-1}\| \leq \epsilon \|\mathcal{H}_j\|$, `Stop`;

`End`

$\mathrm{sgn}(\mathcal{H}) \leftarrow \mathcal{H}_j$;

`Solve` $(I - \mathrm{sgn}(\mathcal{H}))\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 0;$

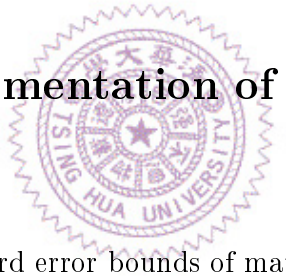
`Compute` $X \leftarrow X_2 X_1^{-1}$.

Remarks:

(i) Notice that $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ spans the stable invariant subspace of the $\mathcal{H}$.

(ii) Both the SDA and the matrix sign function algorithm require $\frac{32}{3}n^3$ flops for each iterative step.

(iii) When working with the Hamiltonian matrix $\mathcal{H}$, a more efficient and structure-preserving version of the classical matrix sign function iteration can be derived by working only with symmetric matrices $J\mathcal{H}$. Details may be consulted in [41, 91].



3 Practical Implementation of SDA

Selection of $\hat{\gamma}$

Here we first derived the forward error bounds of matrices $\widehat{A}_0 \equiv \widehat{A}$, $\widehat{G}_0 \equiv \widehat{G}$ and $\widehat{H}_0 \equiv \widehat{H}$ given in (2.3)–(2.5), respectively. According to these forward errors, we can design a numerical scheme to determine an appropriate value $\hat{\gamma} > 0$. In the following roundoff analysis, we use $fl(\cdot)$ to denote computed floating point values. The quantity $\mathbf{u}$ is the *unit roundoff* (or machine precision), which is typically of order 10^{-8} or 10^{-16} in single and double precision computer arithmetic, respectively. When A and B are $m \times n$ real matrices, the matrix $B := |A|$ if $b_{ij} = |a_{ij}|$ for all i, j, and $A \leq B$ if $a_{ij} \leq b_{ij}$ for all i, j. The 1-, ∞- and Frobenius matrix norms are denoted by $\|\cdot\|_1$, $\|\cdot\|_\infty$ and $\|\cdot\|_F$, respectively.

We assume that the LU factorizations of A_γ and $W_\gamma \equiv A_\gamma + GA_\gamma^{-T}H$ are computed by Gaussian elimination with partial pivoting (GEPP). We write these computed LU factors

as L_A, U_A, L_W and U_W, respectively. Recall that

$$A_\gamma + \Delta A_\gamma = L_A U_A, \quad |\Delta A_\gamma| \le \gamma_n |L_A|\,|U_A|, \tag{3.1}$$

$$W_\gamma + \Delta W_\gamma = L_W U_W, \quad |\Delta W_\gamma| \le \gamma_n |L_W|\,|U_W|, \tag{3.2}$$

with $\gamma_n := n\mathbf{u}/(1-n\mathbf{u})$ (see, e.g., [63, Theorem 9.3]). Then we have

$$fl(W_\gamma^{-1}) = W_\gamma^{-1} + E_1, \quad |E_1| \le c_n\mathbf{u}\,|W_\gamma^{-1}|\,|L_W|\,|U_W|\,|fl(W_\gamma^{-1})|, \tag{3.3}$$

where c_n is a modest constant. From (3.3), the forward error bound in evaluating $\widehat{A}$ in (2.3) is

$$fl(\widehat{A}) = \widehat{A} + E_2, \quad |E_2| \le 4\gamma c_n\mathbf{u}\,|W_\gamma^{-1}|\,|L_W|\,|U_W|\,|fl(W_\gamma^{-1})| + \mathbf{u}\,|\widehat{A}| + O(\mathbf{u^2}). \tag{3.4}$$

Furthermore, from (3.1), we have

$$\widehat{G}_\gamma \equiv fl(2\gamma A_\gamma^{-1}G) = 2\gamma A_\gamma^{-1}G + E_3, \quad |E_3| \le 2\gamma c_n\mathbf{u}\,|A_\gamma^{-1}|\,|L_A|\,|U_A|\,|\widehat{G}_\gamma|, \tag{3.5}$$

hence the forward error bound in evaluating $\widehat{G}$ in (2.4) is

$$\begin{aligned} fl(\widehat{G}) = \widehat{G} + E_4, \quad |E_4| &\le 2\gamma c_n\mathbf{u}\,|A_\gamma^{-1}|\,|L_A|\,|U_A|\,|\widehat{G}_\gamma|\,|W_\gamma^{-1}|^T \\ &\quad + c_n\mathbf{u}\,|fl(\widehat{G})|\,|U_W^T|\,|L_W^T|\,|W_\gamma^{-1}|^T. \end{aligned} \tag{3.6}$$

Finally, from (3.1), we have

$$\widehat{H}_\gamma \equiv fl(2\gamma H A_\gamma^{-1}) = 2\gamma H A_\gamma^{-1} + E_5, \quad |E_5| \le 2\gamma c_n\mathbf{u}\,|\widehat{H}_\gamma|\,|L_A|\,|U_A|\,|A_\gamma^{-1}|, \tag{3.7}$$

and the forward error bound in evaluating $\widehat{H}$ in (2.5) is

$$\begin{aligned} fl(\widehat{H}) = \widehat{H} + E_6, \quad |E_6| &\le 2\gamma c_n\mathbf{u}\,|W_\gamma^{-1}|^T|\widehat{H}_\gamma|\,|L_A|\,|U_A|\,|A_\gamma^{-1}| \\ &\quad + c_n\mathbf{u}\,|W_\gamma^{-1}|^T|U_W^T|\,|L_W^T|\,|fl(\widehat{H})|. \end{aligned} \tag{3.8}$$

For GEPP, we have in practice $\|\,|L_A|\,|U_A|\,\|_\infty \approx \|A_\gamma\|_\infty$ and $\|\,|L_W|\,|U_W|\,\|_\infty \approx \|W_\gamma\|_\infty$, and it follows from (3.4), (3.6) and (3.8) that

$$\|fl(\widehat{A}) - \widehat{A}\|_\infty \lesssim 4c_n\mathbf{u}\gamma\kappa_\infty(W_\gamma)\,\|fl(W_\gamma^{-1})\|_\infty + \mathbf{u}\,\|\widehat{A}\|_\infty + O(\mathbf{u}^2), \tag{3.9}$$

$$\|fl(\widehat{G}) - \widehat{G}\|_\infty \lesssim 2c_n\mathbf{u}\gamma\kappa_\infty(A_\gamma)\,\|W_\gamma^{-1}\|_1\|\widehat{G}_\gamma\|_\infty + c_n\mathbf{u}\kappa_1(W_\gamma)\,\|fl(\widehat{G})\|_\infty, \tag{3.10}$$

$$\|fl(\widehat{H}) - \widehat{H}\|_\infty \le 2c_n\mathbf{u}\gamma\kappa_\infty(A_\gamma)\,\|W_\gamma^{-1}\|_1\|\widehat{H}_\gamma\|_\infty + c_n\mathbf{u}\kappa_1(W_\gamma)\,\|fl(\widehat{H})\|_\infty, \tag{3.11}$$

where $\kappa_1(W_\gamma) \equiv \|W_\gamma\|_1\|W_\gamma^{-1}\|_1$, $\kappa_\infty(W_\gamma) \equiv \|W_\gamma\|_\infty\|W_\gamma^{-1}\|_\infty$ and $\kappa_\infty(A_\gamma) \equiv \|A_\gamma\|_\infty\|A_\gamma^{-1}\|_\infty$.

In order to control the forward error bounds of $\widehat{A}$, $\widehat{G}$ and $\widehat{H}$, we consider the following min-max optimization problem, to determine an optimal value $\hat{\gamma} > 0$:

$$\min_{\gamma>0} F(\gamma) \equiv \max_{i=1,2,3}\{f_i(\gamma)\}, \tag{3.12}$$

where $f_1(\gamma) := \gamma\kappa_\infty(W_\gamma)$, $f_2(\gamma) := \gamma\kappa_\infty(A_\gamma)$ and $f_3(\gamma) := \kappa_1(W_\gamma)$. Since the condition numbers $\kappa_\infty(W_\gamma)$, $\kappa_\infty(A_\gamma)$ and $\kappa_1(W_\gamma)$ approach 1 as $\gamma \to \infty$, it follows that $F(\gamma)$ becomes unbounded as $\gamma \to \infty$. Extensive numerical experiments on randomly generated matrices indicate that $F(\gamma)$ is a strictly convex function in the neighborhood of the optimal $\hat{\gamma}$ where the global minimum of $F(\gamma)$ occurs. For illustration, we report a sample of graphs of $f_1(\gamma)$, $f_2(\gamma)$, $f_3(\gamma)$ and $F(\gamma)$ in Figures 1 and 2. From Theorem 2.1, we know that if γ approaches 0 and ∞, the symplectic matrix pair $(\widehat{\mathcal{N}}, \widehat{\mathcal{L}})$ has eigenvalues close to one, leading to very slow convergence of the SDA. This can be avoided through the min-max optimization problem (3.12).

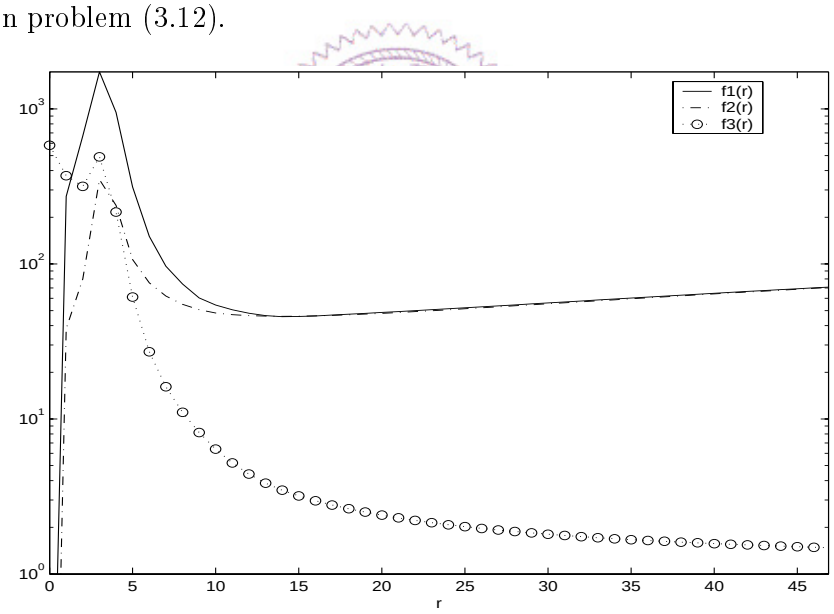


Figure 1: The graphs of functions f_1, f_2 and f_3.

We can apply the Fibonacci search method to compute an approximate value of $\hat{\gamma}$, see, e.g., [12, p. 272]. Our experience indicates that three to five iterations of Fibonacci

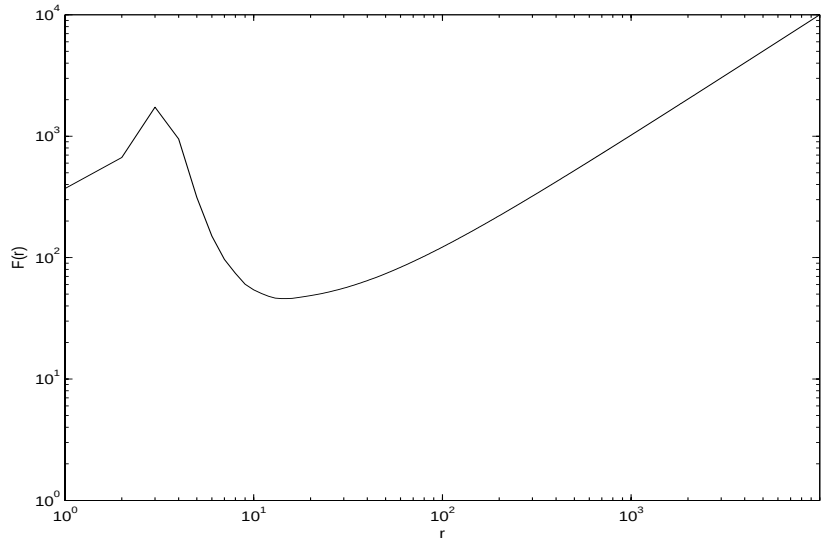


Figure 2: The graph of $F(\gamma)$.

search are adequate to obtain a suboptimal yet acceptable approximation to $\hat{\gamma}$.

Symmetry of $\widehat{G}_0$ and $\widehat{H}_0$

If the matrices G and H are of low-rank, say $G = gg^T \geq 0$ and $H = h^T h \geq 0$, then so are $\widehat{G}_0$ and $\widehat{H}_0$. Indeed, by using the Sherman-Morrison-Woodbury formula (SMWF) twice, it can be seen that

$$\begin{aligned}
\widehat{G}_0 &= 2\gamma A_\gamma^{-1} G (A_\gamma^T + H A_\gamma^{-1} G)^{-1} \\
&= 2\gamma A_\gamma^{-1} g g^T (A_\gamma^T + h h^T A_\gamma^{-1} g g^T)^{-1} \\
&= 2\gamma \left[A_\gamma^{-1} g g^T (A_\gamma^{-T} - A_\gamma^{-T} h h^T (I + A_\gamma^{-1} g g^T A_\gamma^{-T} h h^T)^{-1} A_\gamma^{-1} g g^T A_\gamma^{-T}) \right] \\
&= 2\gamma \left\{ (A_\gamma^{-1} g) \left[I - (A_\gamma^{-1} g)^T h h^T \left(I + (A_\gamma^{-1} g)(A_\gamma^{-1} g)^T h h^T \right)^{-1} (A_\gamma^{-1} g) \right] (A_\gamma^{-1} g)^T \right\} \\
&= 2\gamma \left\{ (A_\gamma^{-1} g) \left(I + (A_\gamma^{-1} g)^T h h^T (A_\gamma^{-1} g) \right)^{-1} (A_\gamma^{-1} g)^T \right\} \\
&= 2\gamma \left\{ (A_\gamma^{-1} g) \left(K_g^T K_g \right)^{-1} (A_\gamma^{-1} g)^T \right\} \quad \text{(Cholesky decomposition)} \\
&= 2\gamma (A_\gamma^{-1} g K_g^{-1})(A_\gamma^{-1} g K_g^{-1})^T.
\end{aligned}$$

Similarly, by applying the same techniques, we also have

$$\begin{aligned}\widehat{H}_0 &= 2\gamma(A_\gamma^T + HA_\gamma^{-1}G)^{-1}HA_\gamma^{-1} \\ &= 2\gamma\left\{(hA_\gamma^{-1})^T\left(I + (hA_\gamma^{-1})gg^T(hA_\gamma^{-1})^T\right)^{-1}(hA_\gamma^{-1})\right\} \\ &= 2\gamma\left\{(hA_\gamma^{-1})^T\left(K_hK_h^T\right)^{-1}(hA_\gamma^{-1})\right\} \quad \text{(Cholesky decomposition)} \\ &= 2\gamma(K_h^{-1}hA_\gamma^{-1})^T(K_h^{-1}hA_\gamma^{-1}).\end{aligned}$$

Computation of $\widehat{A}_j$, $\widehat{G}_j$ and $\widehat{H}_j$

We now propose a structured and efficient procedure for the computation of $\widehat{A}_j$, $\widehat{G}_j$ and $\widehat{H}_j$ in the SDA algorithm, respectively, where $\widehat{G}_0 = \widehat{B}_0\widehat{B}_0^T \geq 0$, $\widehat{H}_0 = \widehat{C}_0^T\widehat{C}_0 \geq 0$ are FRDs. For $j = 0, 1, 2, \ldots$, we let $W_j \equiv (I + \widehat{G}_j\widehat{H}_j)^{-1}$. It is easily seen that $\widehat{H}_jW_j = W_j^T\widehat{H}_j$ and $\widehat{G}_jW_j^T = W_j\widehat{G}_j$ are s.p.s.d. for each $j \geq 1$. By the SMWF we can derive the formulae

$$W_j = (I + \widehat{G}_j\widehat{H}_j)^{-1} = I - \widehat{B}_j(I + \widehat{B}_j^T\widehat{H}_j\widehat{B}_j)^{-1}\widehat{B}_j^T\widehat{H}_j, \tag{3.13}$$

$$\widehat{G}_jW_j^T = \widehat{G}_j - \widehat{G}_j\widehat{C}_j^T(I + \widehat{C}_j\widehat{G}_j\widehat{C}_j^T)^{-1}\widehat{C}_j\widehat{G}_j = \widehat{B}_j(I + \widehat{B}_j^T\widehat{H}_j\widehat{B}_j)^{-1}\widehat{B}_j^T, \tag{3.14}$$

$$W_j^T\widehat{H}_j = \widehat{H}_j - \widehat{H}_j\widehat{B}_j(I + \widehat{B}_j^T\widehat{H}_j\widehat{B}_j)^{-1}\widehat{B}_j^T\widehat{H}_j = \widehat{C}_j^T(I + \widehat{C}_j\widehat{G}_j\widehat{C}_j^T)^{-1}\widehat{C}_j. \tag{3.15}$$

When the matrices B and C start with low ranks in (1.1), we can improve the efficiency of our computation further by the following compression process. Compute the Cholesky decomposition of the s.p.d. matrices $W_{G,j} \equiv (I + \widehat{B}_j^T\widehat{H}_j\widehat{B}_j) = K_{B,j}^TK_{B,j}$ and $W_{H,j} \equiv (I + \widehat{C}_j\widehat{G}_j\widehat{C}_j^T) = K_{C,j}K_{C,j}^T$, respectively. For $j = 0, 1, 2, \ldots$, application of (3.13)–(3.15) leads to

$$\widehat{A}_{j+1} = \widehat{A}_j^2 - \widehat{A}_j\widehat{B}_j(I + \widehat{B}_j^T\widehat{H}_j\widehat{B}_j)^{-1}\widehat{B}_j^T\widehat{H}_j\widehat{A}_j, \tag{3.16}$$

$$\begin{aligned}\widehat{G}_{j+1} &= \widehat{G}_j + \widehat{A}_j\widehat{B}_j(I + \widehat{B}_j^T\widehat{H}_j\widehat{B}_j)^{-1}\widehat{B}_j^T\widehat{A}_j^T \\ &= \begin{bmatrix} \widehat{B}_j, & \widehat{A}_j\widehat{B}_jK_{B,j}^{-1} \end{bmatrix}\begin{bmatrix} \widehat{B}_j^T \\ K_{B,j}^{-T}\widehat{B}_j^T\widehat{A}_j^T \end{bmatrix} \equiv \widehat{B}_{j+1}\widehat{B}_{j+1}^T \geq 0 \quad \text{(FRD)}\end{aligned} \tag{3.17}$$

and

$$\begin{aligned}\widehat{H}_{j+1} &= \widehat{H}_j + \widehat{A}_j^T \widehat{C}_j^T (I + \widehat{C}_j \widehat{G}_j \widehat{C}_j^T)^{-1} \widehat{C}_j \widehat{A}_j \\ &= \begin{bmatrix} \widehat{C}_j^T, & \widehat{A}_j^T \widehat{C}_j^T K_{C,j}^{-T} \end{bmatrix} \begin{bmatrix} \widehat{C}_j \\ K_{C,j}^{-1} \widehat{C}_j \widehat{A}_j \end{bmatrix} \equiv \widehat{C}_{j+1}^T \widehat{C}_{j+1} \geq 0 \quad \text{(FRD)},\end{aligned} \tag{3.18}$$

where $\widehat{B}_{j+1}$ and $\widehat{C}_{j+1}^T$ are the full column rank compressions of $\left[\widehat{B}_j,\ \widehat{A}_j \widehat{B}_j K_{B,j}^{-1}\right]$ and $\left[\widehat{C}_j^T,\ \widehat{A}_j^T \widehat{C}_j^T K_{C,j}^{-T}\right]$, respectively. In general, $\text{rank}(\widehat{B}_{j+1}) > \text{rank}(\widehat{B}_j)$ and $\text{rank}(\widehat{C}_{j+1}) > \text{rank}(\widehat{C}_j)$, and the compression process becomes unprofitable when the ranks of $\widehat{B}_{j+1}$ and $\widehat{C}_{j+1}$ approach n.

Error Analysis of SDA

We consider the forward error bounds of the computed matrices $\widehat{A}_{j+1}$, $\widehat{G}_{j+1}$ and $\widehat{H}_{j+1}$ in the SDA algorithm for one iterative step j. Since $K_{B,j}$ and $K_{C,j}$ are the computed Cholesky factors of matrices $W_{G,j}$ and $W_{H,j}$, respectively, it follows that

$$\begin{aligned}&\widehat{K}_B \equiv fl(K_{B,j}^{-T} \widehat{B}_j^T) = K_{B,j}^{-T} \widehat{B}_j^T + \Delta E_1, \\ &|\Delta E_1| \leq c_1 \mathbf{u}\, |K_{B,j}^{-T}|\, |K_{B,j}^T|\, |\widehat{K}_B|,\end{aligned} \tag{3.19}$$

and

$$\begin{aligned}&\widehat{K}_C \equiv fl(K_{C,j}^{-1} \widehat{C}_j) = K_{C,j}^{-1} \widehat{C}_j + \Delta \widetilde{E}_1, \\ &|\Delta \widetilde{E}_1| \leq \tilde{c}_1 \mathbf{u}\, |K_{C,j}^{-1}|\, |K_{C,j}|\, |\widehat{K}_C|,\end{aligned} \tag{3.20}$$

where c_1 and $\tilde{c}_1$ are modest constants. Therefore, we have

$$\begin{aligned}&fl(K_{B,j}^{-T} \widehat{B}_j^T \widehat{A}_j^T) = fl(\widehat{K}_B \widehat{A}_j^T) = K_{B,j}^{-T} \widehat{B}_j^T \widehat{A}_j^T + \Delta E_2, \\ &|\Delta E_2| \leq c_2 \mathbf{u}\, |K_{B,j}^{-T}|\, |K_{B,j}^T|\, |\widehat{K}_B|\, |\widehat{A}_j^T|,\end{aligned} \tag{3.21}$$

and

$$\begin{aligned}&fl(K_{C,j}^{-1} \widehat{C}_j \widehat{A}_j) = fl(\widehat{K}_C \widehat{A}_j) = K_{C,j}^{-1} \widehat{C}_j \widehat{A}_j + \Delta \widetilde{E}_2, \\ &|\Delta \widetilde{E}_2| \leq \tilde{c}_2 \mathbf{u}\, |K_{C,j}^{-1}|\, |K_{C,j}|\, |\widehat{K}_C|\, |\widehat{A}_j|,\end{aligned} \tag{3.22}$$

where c_2 and $\tilde{c}_2$ are modest constants.

If $\text{rank}(\widehat{B}_j^T) = \ell$, then from Theorem 18.4 of [63] and (3.17), there exist an orthogonal matrix $Q_B \in \mathbb{R}^{2\ell \times 2\ell}$ and a computed upper triangular matrix $fl(\widehat{B}_{j+1}^T)$ with full row rank, such that

$$\begin{bmatrix} \widehat{B}_j^T \\ fl(\widehat{K}_B \widehat{A}_j^T) \end{bmatrix} + \begin{bmatrix} \Delta B_1 \\ \Delta B_2 \end{bmatrix} = Q_B \begin{bmatrix} fl(\widehat{B}_{j+1}^T) \\ 0 \end{bmatrix}, \tag{3.23}$$

where $|\Delta B_j| \leq c_3 \mathbf{u}\ G_\ell(|\widehat{B}_j^T| + |K_{B,j}^{-T}|\ |\widehat{B}_j^T|\ |\widehat{A}_j^T|)$ for $j = 1, 2$, with c_3 being a modest constant and $\|G_\ell\|_F = \frac{1}{2}$.

From (3.21) and (3.23), we have

$$\begin{bmatrix} \widehat{B}_j^T \\ K_{B,j}^{-T} \widehat{B}_j^T \widehat{A}_j^T \end{bmatrix} + \begin{bmatrix} \Delta B_1 \\ \Delta \widetilde{B}_2 \end{bmatrix} = Q_B \begin{bmatrix} fl(\widehat{B}_{j+1}^T) \\ 0 \end{bmatrix}, \tag{3.24}$$

where $|\Delta \widetilde{B}_2| \leq c_2 \mathbf{u}\ |K_{B,j}^{-T}|\ |K_{B,j}^T|\ |\widehat{K}_B|\ |\widehat{A}_j^T| + c_3 \mathbf{u}\ G_\ell(|\widehat{B}_j^T| + |K_{B,j}^{-T}|\ |\widehat{B}_j^T|\ |\widehat{A}_j^T|)$. Pre-multiplying both sides of (3.24) by Q_B^T, it follows that

$$\begin{bmatrix} \widehat{B}_{j+1}^T \\ 0 \end{bmatrix} + Q_B^T \begin{bmatrix} \Delta B_1 \\ \Delta \widetilde{B}_2 \end{bmatrix} = \begin{bmatrix} fl(\widehat{B}_{j+1}^T) \\ 0 \end{bmatrix}, \tag{3.25}$$

and we deduce that

$$\|fl(\widehat{B}_{j+1}^T) - \widehat{B}_{j+1}^T\|_F \leq c_4 \mathbf{u}\ \|\widehat{B}_j\|_F + c_5 \mathbf{u}\ \kappa_s(K_{B,j}^T)\ \|\widehat{K}_b\|_F \|\widehat{A}_j\|_F, \tag{3.26}$$

where c_4 and c_5 are modest constants, and $\kappa_s(K_{B,j}^T) \equiv \| |K_{B,j}^{-T}|\ |K_{B,j}^T| \|_\infty$ is the Skeel condition number of $K_{B,j}^T$. Furthermore, applying a similar argument with the help of (3.22), we can derive that

$$\|fl(\widehat{C}_{j+1}) - \widehat{C}_{j+1}\|_F \leq c_6 \mathbf{u}\ \|\widehat{C}_j\|_F + c_7 \mathbf{u}\ \kappa_s(K_{C,j})\ \|\widehat{K}_C\|_F \|\widehat{A}_j\|_F, \tag{3.27}$$

where c_6 and c_7 are modest constants.

On the other hand, it follows from (3.19) that

$$\begin{aligned} &fl(K_{B,j}^{-T} \widehat{B}_j^T \widehat{H}_j \widehat{A}_j) = K_{B,j}^{-T} \widehat{B}_j^T \widehat{H}_j \widehat{A}_j + \Delta E_3, \\ &|\Delta E_3| \leq c_8 \mathbf{u}\ |K_{B,j}^{-T}|\ |K_{B,j}^T|\ |\widehat{K}_B|\ |\widehat{H}_j|\ |\widehat{A}_j|, \end{aligned} \tag{3.28}$$

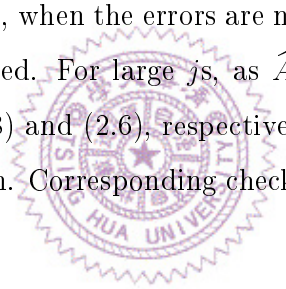
where c_8 is a modest constant. From (3.21) and (3.28), the forward error bound of computing $\widehat{A}_{j+1}$ is

$$\|fl(\widehat{A}_{j+1}) - \widehat{A}_{j+1}\|_F \leq c_9\mathbf{u}\,\|\widehat{A}_j\|_F^2 + c_{10}\mathbf{u}\,\kappa_s(K_{B,j}^T)\,\|\widehat{B}_j\|_F^2\|\widehat{K}_{B,j}^{-1}\|_F^2\|\widehat{H}_j\|_F\|\widehat{A}_j\|_F^2, \quad (3.29)$$

where c_9 and c_{10} are modest constants.

When the Skeel condition numbers $\kappa_s(K_{B,j}^T)$ and $\kappa_s(K_{C,j})$ in (3.26) and (3.27) are bounded from above by acceptable numbers, the accumulation of error will be dampened by the fast rate of convergence at the final stage of the iterative process. Danger, if any, lies in the early stage of the process before the $\lambda_n^{2^j}$ convergence factor dominates. It is unlikely to have any ill-effect, as the accumulated error in the matrix additions and multiplications should be of magnitude around a small multiple of the machine accuracy.

As the SSF properties are preserved in the SDA, any error will be a structured one, only pushing the iteration towards a solution of a neighboring SSF system. Thus the algorithm is stable in this sense, when the errors are not too large and when stabilizability and detectability are maintained. For large js, as $\widehat{A}_j \to 0$, $\widehat{G}_j$ and $\widehat{H}_j$ converge to the unique s.p.s.d. solutions of (2.8) and (2.6), respectively. Danger again will only comes at the initial stage of the iteration. Corresponding checks may be prudent in the algorithm.



4 SDA_m

A matrix A is persymmetric when A is symmetric with respect to the main anti-diagonal ([59, p. 193]). When the DARE transformed from the CARE (1.1) has the additional property that the initial data $\widehat{A}_0$, $\widehat{G}_0 = \widehat{H}_0 \in \mathbb{R}^{2\ell\times 2\ell}$ are symmetric and persymmetric, the additional structure can be preserved in a modified version of the SDA (SDA_m). For simplicity, we consider only when $\gamma = 1$. This doubly symmetric type of DAREs appear in the Examples 10 and 17 of Section 5 (originally from [22]).

For convenience, in the SDA, we denote for $j = 1, 2, \cdots$

$$\begin{aligned} A \equiv \widehat{A}_j, \quad G \equiv \widehat{G}_j = \widehat{H}_j, \\ A_+ \equiv \widehat{A}_{j+1}, \quad G_+ \equiv \widehat{G}_{j+1} = \widehat{H}_{j+1}. \end{aligned} \quad (4.1)$$

Since A, $G = H$ are symmetric and persymmetric of even order, we write

$$A = \begin{bmatrix} a_1 & a_2\zeta \\ \zeta a_2 & \zeta a_1 \zeta \end{bmatrix}, \quad G = \begin{bmatrix} g_1 & g_2\zeta \\ \zeta g_2 & \zeta g_1 \zeta \end{bmatrix}, \tag{4.2}$$

where a_1, a_2, g_1 and $g_2 \in \mathbb{R}^{\ell\times\ell}$ are symmetric and $\zeta = [e_\ell, \cdots, e_1]$ with e_j being the jth column of I_ℓ. In the SDA, we shall show that $\widehat{A}$, $\widehat{G}$ and $\widehat{H}$ are also symmetric and persymmetric with $\widehat{G} = \widehat{H}$, with

$$A_+ = A(I + G^2)^{-1}A = \begin{bmatrix} \widehat{a}_1 & \widehat{a}_2\zeta \\ \zeta\widehat{a}_2 & \zeta\widehat{a}_1\zeta \end{bmatrix}, \tag{4.3}$$

$$G_+ = G + AG(I + G^2)^{-1}A^T = \begin{bmatrix} \widehat{g}_1 & \widehat{g}_2\zeta \\ \zeta\widehat{g}_2 & \zeta\widehat{g}_1\zeta \end{bmatrix}. \tag{4.4}$$

Let

$$q_1 \equiv g_1 + g_2, \quad q_2 \equiv g_1 - g_2, \quad \alpha_1 \equiv a_1 + a_2 \quad , \quad \alpha_2 \equiv a_1 - a_2. \tag{4.5}$$

Simple manipulation leads to

$$\widehat{a}_1 = \frac{1}{2}\left[\alpha_1(I + q_1^2)^{-1}\alpha_1 + \alpha_2(I + q_2^2)^{-1}\alpha_2\right], \tag{4.6}$$

$$\widehat{a}_2 = \frac{1}{2}\left[\alpha_1(I + q_1^2)^{-1}\alpha_1 - \alpha_2(I + q_2^2)^{-1}\alpha_2\right], \tag{4.7}$$

$$\widehat{g}_1 = g_1 + \frac{1}{2}\left[\alpha_1(I + q_1^2)^{-1}q_1\alpha_1 + \alpha_2(I + q_2^2)^{-1}q_2\alpha_2\right], \tag{4.8}$$

$$\widehat{g}_2 = g_2 + \frac{1}{2}\left[\alpha_1(I + q_1^2)^{-1}q_1\alpha_1 - \alpha_2(I + q_2^2)^{-1}q_2\alpha_2\right]. \tag{4.9}$$

Furthermore let

$$q_1 = [U_1, V_1]\begin{bmatrix} \Sigma_1 & 0 \\ 0 & -\Gamma_1 \end{bmatrix}\begin{bmatrix} U_1^T \\ V_1^T \end{bmatrix}, \quad q_2 = [U_2, V_2]\begin{bmatrix} \Sigma_2 & 0 \\ 0 & -\Gamma_2 \end{bmatrix}\begin{bmatrix} U_2^T \\ V_2^T \end{bmatrix} \tag{4.10}$$

be the spectral decompositions of q_1 and q_2, respectively, with Σ_1, Γ_1, Σ_2 and Γ_2 being nonnegative diagonal matrices. Then $\widehat{a}_1$, $\widehat{a}_2$, $\widehat{g}_1$ and $\widehat{g}_2$ in (4.6)–(4.9) can be computed by

the following symmetric forms:

$$\begin{aligned}
\xi_1 &\equiv \alpha_1 U_1 (I + \Sigma_1^2)^{-1} U_1^T \alpha_1 - \alpha_1 V_1 (I + \Gamma_1^2)^{-1} V_1^T \alpha_1, \\
\xi_2 &\equiv \alpha_2 U_2 (I + \Sigma_2^2)^{-1} U_2^T \alpha_2 - \alpha_2 V_2 (I + \Gamma_2^2)^{-1} V_2^T \alpha_2, \\
\widehat{a}_1 &= \frac{1}{2}\{\xi_1 + \xi_2\}, \quad \widehat{a}_2 = \frac{1}{2}\{\xi_1 - \xi_2\}; \qquad (4.11) \\
\eta_1 &\equiv \alpha_1 U_1 (I + \Sigma_1^2)^{-1} \Sigma_1 U_1^T \alpha_1 - \alpha_1 V_1 (I + \Gamma_1^2)^{-1} \Gamma_1 V_1^T \alpha_1, \\
\eta_2 &\equiv \alpha_2 U_2 (I + \Sigma_2^2)^{-1} \Sigma_2 U_2^T \alpha_2 - \alpha_2 V_2 (I + \Gamma_2^2)^{-1} \Gamma_2 V_2^T \alpha_2, \\
\widehat{g}_1 &= g_1 + \frac{1}{2}\{\eta_1 + \eta_2\}, \quad \widehat{g}_2 = g_2 + \frac{1}{2}\{\eta_1 - \eta_2\}. \qquad (4.12)
\end{aligned}$$

The SDA_m computes $\widehat{A}$, $\widehat{G}$ in (4.3) and (4.4) using the symmetric forms (4.11) and (4.12) and considerably improves the accuracy of Examples 10 and 17 in the next section.

5 Numerical Examples

For the Tables in the following examples, data for various methods are lists in columns with obvious headings. The heading "`care`" is for the `care` command in MATLAB [88], "MSGM" is for the matrix sign function method [41], and "SDA" (or "SDA_m") stands for our SDA (or SDA_m) algorithm. There is no iteration numbers to report for `care` and an '$*$' in the Tables indicates a failure of convergence in obtaining a solution. In the graphs, "ratio_care" and "ratio_MSGM" are the ratio of the CPU-times for `care` and MSGM to that of the SDA, respectively. For the comparison of residuals, the "normalized" residual (NRes) formula is applied in the numerical examples, i.e.,

$$\text{NRes} \equiv \frac{\|A^T\tilde{X} + \tilde{X}A^T - \tilde{X}G\tilde{X}A + H\|}{\|A^T\tilde{X}\| + \|\tilde{X}A^T\| + \|\tilde{X}G\tilde{X}\| + \|H\|},$$

where $\tilde{X}$ is an approximate solution and $\|\cdot\|$ denotes the 2-norm for matrices.

Some numerical examples from [22] involved very large data sets, which have not been repeated here. Twenty examples were presented in [45]. We retain the numbering of examples in [45], comment upon all of them but present only five representative ones in this chapter.

In the MSGM, the scaling strategy suggested in [41] was implemented. For a fairer comparison, similar convergence criteria were used in all the methods and the solutions were not refined.

All computations were performed using MATLAB/Version 6.0 on a Compaq/DS20 workstation. The machine precision is 2.22×10^{-16}.

Example 5. The example is identical to Example 5 of [22], which has been presented originally in [97]. This is a 9th-order continuous state space model of a tubular ammonia reactor. The actual system matrices are

$$A = \begin{bmatrix} -4.019 & 5.12 & 0 & 0 & -2.082 & 0 & 0 & 0 & 0.87 \\ -0.346 & 0.986 & 0 & 0 & -2.34 & 0 & 0 & 0 & 0.97 \\ -7.909 & 15.407 & -4.096 & 0 & -6.45 & 0 & 0 & 0 & 2.68 \\ -21.816 & 35.606 & -0.339 & -3.87 & -17.8 & 0 & 0 & 0 & 7.39 \\ -60.196 & 98.188 & -7.907 & 0.34 & -53.008 & 0 & 0 & 0 & 20.4 \\ 0 & 0 & 0 & 0 & 94.0 & -147.2 & 0 & 53.2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 94.0 & -147.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 12.8 & 0 & -31.6 & 0 \\ 0 & 0 & 0 & 0 & 12.8 & 0 & 0 & 18.8 & -31.6 \end{bmatrix},$$

$$B^T = \begin{bmatrix} 0.010 & 0.003 & 0.009 & 0.024 & 0.068 & 0 & 0 & 0 & 0 \\ -0.011 & -0.021 & -0.059 & -0.162 & -0.445 & 0 & 0 & 0 & 0 \\ -0.151 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad H = I_9, \quad R = I_3 .$$

The numerical results are given in Table 1.

	SDA	MSGM	`care`
NRes	1.68×10^{-15}	1.73×10^{-13}	4.64×10^{-14}
Iter. no.	9	8	-

Table 1: Results for Example 5.

Example 6. The example is identical to Example 6 of [22], which has been presented

originally in [49]. This control problem for a J-100 jet engine is a special case of a multivariable servomechanism problem. To save space, we shall not list the system matrices here. We report the numerical results in Table 2.

	SDA	MSGM	`care`
NRes	5.78×10^{-13}	3.11×10^{-8}	1.91×10^{-12}
Iter. no.	10	9	-

Table 2: Results for Example 6.

Example 10. The example is identical to Example 10 of [22], which has been presented originally in [8]. Here, the system matrices are

$$A = \begin{bmatrix} \varepsilon + 1 & 1 \\ 1 & \varepsilon + 1 \end{bmatrix}, \quad G = I_2, \quad H = \begin{bmatrix} \varepsilon^2 & 0 \\ 0 & \varepsilon^2 \end{bmatrix}.$$

The exact stabilizing solution X is given by

$$x_{11} = x_{22} = \frac{1}{2}\left[2(\varepsilon+1) + \sqrt{2(\varepsilon+1)^2 + 2 + \sqrt{2}\varepsilon}\right], \quad x_{12} = x_{21} = x_{11}/[x_{11} - (\varepsilon+1)] .$$

The corresponding DARE is doubly symmetric and the SDA_m was applied (see details in Section 4). The numerical results with $\varepsilon = 1, 10^{-3}, 10^{-5}$ and 10^{-7} are given in Table 3.

Example 11. The example is identical to Example 11 of [22], which has been presented originally in [66]. This example represents an algebraic Riccati equation arising from a H_∞-control problem [131]. Let

$$A = \begin{bmatrix} 3 - \varepsilon & 1 \\ 4 & 2 - \varepsilon \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad R = 1, \quad H = \begin{bmatrix} 4\varepsilon - 11 & 2\varepsilon - 5 \\ 2\varepsilon - 5 & 2\varepsilon - 2 \end{bmatrix}.$$

The matrix

$$X = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

		SDA	SDA_m	MSGM	`care`
$\varepsilon = 1$	NRes	0.00×10^{0}	0.00×10^{0}	4.69×10^{-16}	9.36×10^{-17}
	Rel. err.	1.96×10^{-16}	1.96×10^{-16}	8.80×10^{-16}	3.83×10^{-16}
	Iter. no.	4	4	2	-
$\varepsilon = 10^{-3}$	NRes	1.58×10^{-14}	1.43×10^{-16}	1.11×10^{-13}	9.20×10^{-17}
	Rel. err.	1.82×10^{-11}	2.22×10^{-16}	2.22×10^{-13}	4.08×10^{-16}
	Iter. no.	16	13	12	-
$\varepsilon = 10^{-5}$	NRes	2.28×10^{-12}	1.11×10^{-16}	1.07×10^{-11}	5.53×10^{-17}
	Rel. err.	7.16×10^{-7}	1.76×10^{-16}	2.14×10^{-11}	2.60×10^{-16}
	Iter. no.	22	19	18	-
$\varepsilon = 10^{-7}$	NRes	1.49×10^{-10}	1.32×10^{-16}	3.31×10^{-9}	2.06×10^{-17}
	Rel. err.	6.04×10^{-8}	4.44×10^{-16}	6.63×10^{-9}	1.36×10^{-16}
	Iter. no.	12	26	20	-

Table 3: Results for Example 10.

is the stabilizing solution for $\varepsilon > 0$. For $\varepsilon = 0$, the solution X is obtained by an H-invariant *Lagrangian* subspace, i.e., a solution in the sense of H_∞-control. The numerical results with $\varepsilon = 1, 0$ are given in Table 4.

Example 12. The example is identical to Example 12 of [22], which has been presented originally in [65]. Let

$$V = I - \frac{2}{3}vv^T, \quad v^T = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \; ; \quad A_0 = \varepsilon \, \mathrm{diag}(1, 2, 3), \quad H_0 = \mathrm{diag}(\varepsilon^{-1}, 1, \varepsilon) \; ;$$

we have

$$A = VA_0V, \quad G = \varepsilon^{-1} I_3, \quad H = VH_0V.$$

The solution is

$$X = V \, \mathrm{diag}(x_1, x_2, x_3) \, V$$

		SDA	MSGM	`care`
$\varepsilon = 1$	NRes	0.00×10^{0}	1.69×10^{-16}	1.97×10^{-16}
	Rel. err.	1.26×10^{-16}	1.25×10^{-15}	9.68×10^{-16}
	Iter. no.	5	2	-
$\varepsilon = 0$	NRes	3.06×10^{-16}	$*$	5.06×10^{-17}
	Rel. err.	2.66×10^{-9}	$*$	7.68×10^{-9}
	Iter. no.	28	$*$	-

Table 4: Results for Example 11.

where

$$x_1 = \varepsilon^2 + \sqrt{\varepsilon^4 + 1}, \quad x_2 = 2\varepsilon^2 + \sqrt{4\varepsilon^4 + \varepsilon}, \quad x_3 = 3\varepsilon^2 + \sqrt{9\varepsilon^4 + \varepsilon^2}.$$

The numerical results with $\varepsilon = 1, 10^6$ are given in Table 5.

		SDA	MSGM	`care`
$\varepsilon = 1$	NRes	2.01×10^{-16}	1.78×10^{-15}	3.00×10^{-16}
	Rel. err.	4.33×10^{-16}	2.78×10^{-15}	5.03×10^{-16}
	Iter. no.	6	4	-
$\varepsilon = 10^6$	NRes	1.62×10^{-15}	2.22×10^{-4}	2.19×10^{-15}
	Rel. err.	2.58×10^{-15}	6.33×10^{-4}	4.92×10^{-15}
	Iter. no.	11	10	-

Table 5: Results for Example 12.

Example 15. The example is identical to Example 15 of [22], which has been presented originally in [75, Example 4] and [4]. This example arises from a mathematical model of position and velocity control for a string of high-speed vehicles. If N vehicles are to be controlled, the size of the system matrices will be $n = 2N - 1$, the number of control inputs will be $m = N$, and the number of outputs will be $p = N - 1$, respectively. The

comparison of normalized residuals are reported in Table 6 for $N = 5, 20, 60, 100, 140$ and 180. Figure 3 reports the comparison of CPU times for `care`, MSGN and the SDA.

		SDA	MSGM	`care`
$N = 5$	NRes	1.61×10^{-16}	8.75×10^{-15}	2.53×10^{-15}
	Iter. no.	5	6	-
$N = 20$	NRes	3.85×10^{-16}	3.55×10^{-14}	6.15×10^{-15}
	Iter. no.	5	6	-
$N = 60$	NRes	1.53×10^{-15}	2.32×10^{-13}	8.14×10^{-15}
	Iter. no.	7	8	-
$N = 100$	NRes	2.15×10^{-15}	6.62×10^{-13}	2.55×10^{-14}
	Iter. no.	8	9	-
$N = 140$	NRes	3.05×10^{-15}	6.50×10^{-12}	3.60×10^{-14}
	Iter. no.	8	9	-
$N = 180$	NRes	1.25×10^{-14}	4.64×10^{-12}	2.01×10^{-13}
	Iter. no.	9	9	-

Table 6: Comparison of normalized residuals for Example 15.

Example 17. The example is identical to Example 17 of [22], which has been presented originally in [75, Example 6]. The system matrices are

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad R = r, \quad C^T = \sqrt{q} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}.$$

It is known from [75] that $x_{1n} = \sqrt{qr}$. Therefore, we may use the relative error in x_{1n}, i.e., $\text{RE} \equiv (|x_{1n} - \sqrt{qr}|)/\sqrt{qr}$, as an indicator of the accuracy of the results. The corresponding

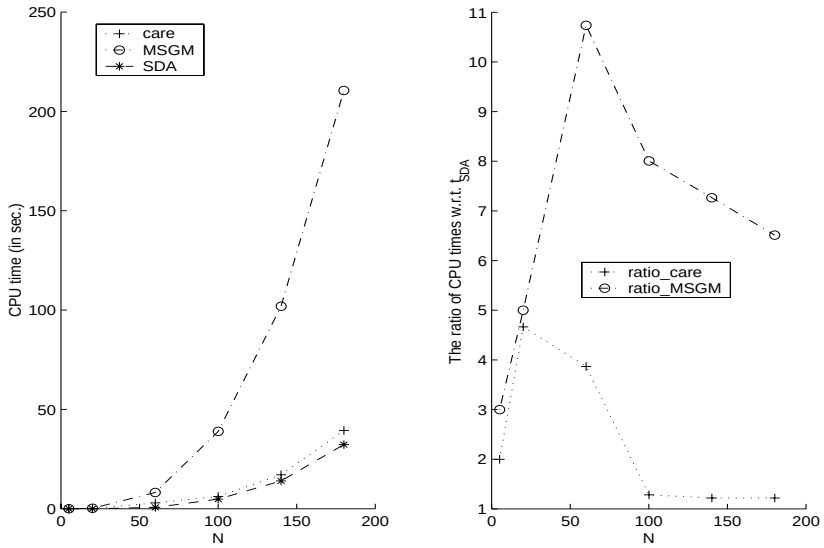


Figure 3: Comparison of CPU times for Example 15.

DARE is doubly symmetric and the SDA_m was applied (see details in Section 4). Table 7 reports the comparison of normalized residuals computed by SDA, SDA_m and `care` for $n = 6, 12, 18, 24, 30$. We also report the comparison of relative errors in x_{1n} computed by above three algorithms in Table 8.

Comments on Numerical Results

We have tested twenty examples in [45] to illustrate the accuracy and efficiency of the SDA applied to CAREs, in comparison to the MSGM [41] and `care` in MATLAB [88]. Some of these examples have parameters to vary their sizes or conditioning. In what follows, we shall comment upon all the examples in [45], thus retaining the old labelling of the examples:

(1) Comparing with `care` for all the examples, solutions with better or comparable accuracy were obtained using the SDA in far less time. This comparison has been difficult as `care` yields no iteration numbers and the CPU time information from MATLAB is not always accurate.

	n	NRes_SDA	NRes_SDA_m	NRes_MSGM	NRes_care
$q, r = 1$	6	4.50×10^{-15}	3.56×10^{-16}	8.87×10^{-15}	1.80×10^{-14}
	12	3.63×10^{-10}	3.22×10^{-14}	9.68×10^{-12}	1.23×10^{-11}
	18	9.47×10^{-5}	1.83×10^{-11}	4.63×10^{-9}	9.46×10^{-9}
	24	2.47×10^{-2}	2.34×10^{-8}	9.88×10^{-6}	3.25×10^{-7}
	30	4.80×10^{-1}	3.52×10^{-5}	3.47×10^{-2}	7.17×10^{-4}
$q, r = 100$	6	2.59×10^{-15}	2.82×10^{-16}	1.20×10^{-11}	1.02×10^{-15}
	12	4.81×10^{-10}	2.94×10^{-14}	4.11×10^{-9}	1.58×10^{-11}
	18	4.33×10^{-5}	2.26×10^{-11}	1.78×10^{-6}	7.83×10^{-9}
	24	7.24×10^{-1}	2.90×10^{-8}	1.37×10^{-2}	1.50×10^{-5}
	30	3.07×10^{-1}	1.45×10^{-5}	2.94×10^{-1}	4.38×10^{-3}

Table 7: Comparison of normalized residuals for Example 17.

(2) The best indication of the efficiency of the SDA over `care` comes from Example 15 (with varying dimension n), where `care` required two to eight times more CPU times than the SDA. This is consistent with the findings in Chapter 1 for DAREs. Keep in mind that the SDA requires far less number of flops than `care` in each iteration, as the operations in the SDA are performed in $\Re^{n\times n}$ whereas those for `care` are carried out in $\Re^{2n\times 2n}$.

(3) For examples with varying conditioning, such as Examples 9-14, 17 and 18, the SDA out-performed `care` and converges to more accurate solutions in less time. For the ill-conditioned Example 20, `care` failed while the SDA succeeded without difficulty.

(4) In Example 11 (in H_∞ control), some eigenvalues were numerically on the imaginary axis and assumptions in the theory were practically violated. The stronger structure-preserving property of the SDA enabled it to produce an accurate solution when the MSGM failed. Somehow, `care` produced a slightly less accurate solution using much more CPU time.

	n	RE_SDA	RE_SDA_m	RE_MSGM	RE_care
$q, r = 1$	6	1.94×10^{-14}	1.11×10^{-15}	2.22×10^{-14}	5.68×10^{-14}
	12	3.99×10^{-10}	1.68×10^{-13}	4.92×10^{-11}	5.61×10^{-11}
	18	8.73×10^{-5}	6.37×10^{-11}	2.35×10^{-8}	2.68×10^{-8}
	24	3.09×10^{-1}	6.39×10^{-8}	3.29×10^{-5}	6.16×10^{-6}
	30	6.49×10^{-1}	1.57×10^{-4}	1.40×10^{-1}	8.37×10^{-3}
$q, r = 100$	6	2.13×10^{-15}	9.95×10^{-16}	3.27×10^{-11}	7.67×10^{-15}
	12	6.04×10^{-10}	1.83×10^{-13}	2.30×10^{-8}	6.13×10^{-11}
	18	5.87×10^{-4}	1.16×10^{-10}	2.95×10^{-5}	2.70×10^{-8}
	24	2.02×10^{-1}	1.32×10^{-7}	2.39×10^{-2}	5.20×10^{-5}
	30	4.60×10^{-1}	5.67×10^{-5}	2.98×10^{-1}	1.71×10^{-2}

Table 8: Comparison of relative errors in x_{1n} for Example 17.

(5) In Examples 10 and 17, the CAREs gave rise to DAREs which were "doubly symmetric" (see Section 4 for details). The SDA_m improved the efficiency of the SDA for these examples, obtaining comparable accuracy for Example 10 while outperforming `care` for Example 17.

(6) Comparing to the MSGM for ill-conditioned problems, the SDA performed better in terms of accuracies or number of iterations. This is consistent with the fact that while both the SDA and MSGM are structure-preserving, the former preserves more structure than the latter. For some well-conditioned problems, the efficiency and accuracy of the SDA and MSGM are comparable. For a few simple small examples, the MSGM converged quickly and was superior to the SDA. Note that the work involved in an iteration for either method is similar.

(7) The MSGM, with similar operations count to SDA, was generally more efficient than `care`, especially for well-conditioned problems. For ill-conditioned problems (such as Example 10), the MSGM was sometimes less accurate than `care`.

6 Conclusions

Solving CAREs as DAREs, after applying the Cayley transform, has previously been investigated by many. Recent developments and better understanding of doubling algorithms, especially the structure-preserving properties and efficiency of the SDA proposed in Chapter 1, give this old approach a new lease of life. In addition, we have studied how the parameter γ in the Cayley transform can be chosen optimally. A Fibonacci search for choosing γ was suggested in Section 3, together with the details of other issues involved in the practical implementation of the SDA. We have also developed the SDA_m which preserves the structure of some doubly symmetric DAREs. Extensive numerical results show that this approach of solving CAREs using the SDA is efficient and competitive, especially for ill-conditioned problems.

Chapter 3

Structure-Preserving Doubling Algorithm for G-DAREs

1 Introduction

Let matrices E, $A \in \mathbb{R}^{n\times n}$ with E being nonsingular, $Q \in \mathbb{R}^{p\times p}$ with $Q = Q^T > 0$ (symmetric positive definite or s.p.d.), $B \in \mathbb{R}^{n\times m}$, $S \in \mathbb{R}^{m\times p}$ and $C \in \mathbb{R}^{n\times p}$ with B, C possessing full column rank, and the s.p.d. $R \in \mathbb{R}^{m\times m}$. Suppose further that the matrix $Q - S^T R^{-1} S$ is symmetric positive semidefinite (s.p.s.d.). The generalized discrete-time algebraic Riccati equation (G-DARE) has the form

$$E^T X E = A^T X A + (A^T X B + C S^T)(R + B^T X B)^{-1}(B^T X A + S C^T) + C Q C^T. \quad (1.1)$$

Equation (1.1) arises frequently in discrete-time optimal control problems and optimal filter problems [75, 76, 77, 91] for a given descriptor linear system:

$$\begin{cases} E x_{k+1} = A x_k + B u_k, \quad x_0 = x^0 \\ y_k = C^T x_k \end{cases} \quad (1.2)$$

with the control vectors $\{u_k\}$ chosen through

$$\min_{u_k} J \equiv \frac{1}{2} \sum_{k=0}^{\infty} (y_k^T Q y_k + u_k^T R u_k + y_k^T S^T u_k + u_k^T S y_k). \quad (1.3)$$

Let

$$C(Q - S^T R^{-1} S) C^T = \widehat{C}_0 \widehat{C}_0^T \geq 0 \quad (1.4)$$

be in full rank decomposition (FRD). The systems denoted by (E, A, B) and $(E, A, \widehat{C}_0)$ are assumed to be stabilizable (S) and detectable (D), respectively. Note that (E, A, B) is stabilizable if $w^T B = 0$ and $w^T A = \lambda E w^T$ imply $|\lambda| < 1$ or $w = 0$. The system $(E, A, \widehat{C}_0)$

is detectable if $(E^T, A^T, \widehat{C}_0^T)$ is stabilizable. The optimal feedback control $\{u_k^*\}$ for (1.2) and (1.3) are given by

$$u_k^* = -(R + B^T X_+ B)^{-1}(B^T X_+ A + SC^T)x_k, \tag{1.5}$$

where $X_+ \geq 0$ is the unique s.p.s.d. solution to (1.1). Furthermore, the closed-loop dynamics of the system obtained with this control

$$Ex_{k+1} = (A + BK)x_k = \left[A - B(R + B^T X_+ B)^{-1}(B^T X_+ A + SC^T)\right] x_k \tag{1.6}$$

is asymptotically stable, i.e., $\lim_{k\to\infty} x_k = 0$ (see, e.g., [91]).

It is well-known [91] that the s.p.s.d. solution of the G-DARE (1.1) can be obtained via the computation of the stable deflating subspace of the matrix pencil

$$\mathcal{A} - \lambda\mathcal{B} = \begin{bmatrix} A & 0 & B \\ -CQC^T & E^T & -CS^T \\ SC^T & 0 & R \end{bmatrix} - \lambda \begin{bmatrix} E & 0 & 0 \\ 0 & A^T & 0 \\ 0 & -B^T & 0 \end{bmatrix}. \tag{1.7}$$

If the columns of $[I_n, E^T X_+, Z]^T$ span the stable deflating subspace of $\mathcal{A} - \lambda\mathcal{B}$, then $X_+ \geq 0$ solves the G-DARE (1.1). Here I_n denotes the identity matrix of order n.

It is also well-known [91] that the s.p.s.d. solution X_+ of G-DARE can be solved by computing the stable deflating subspace $\text{span}\left\{\begin{bmatrix} I_n \\ X_+E \end{bmatrix}\right\}$ of the reduced matrix pencil of (1.7):

$$\mathcal{M}_0 - \lambda\mathcal{L}_0 = \begin{bmatrix} A - BR^{-1}SC^T & 0 \\ -C(Q - S^T R^{-1} S)C^T & E^T \end{bmatrix} - \lambda \begin{bmatrix} E & BR^{-1}B^T \\ 0 & A^T - CS^T R^{-1} B^T \end{bmatrix}. \tag{1.8}$$

Furthermore, it is easily seen that the pencil $\mathcal{M}_0 - \lambda\mathcal{L}_0$ is equivalent to the symplectic pencil

$$\mathcal{M} - \lambda\mathcal{L} = \begin{bmatrix} E^{-1}A - E^{-1}BR^{-1}SC^T & 0 \\ -C(Q - S^T R^{-1} S)C^T & I \end{bmatrix} - \lambda \begin{bmatrix} I & E^{-1}BR^{-1}B^T E^{-T} \\ 0 & A^T E^{-T} - CS^T R^{-1} B^T E^{-T} \end{bmatrix}. \tag{1.9}$$

If the column of $\begin{bmatrix} I \\ X_* \end{bmatrix}$ span the stable deflating subspace of $\mathcal{M} - \lambda\mathcal{L}$, then $E^{-T} X_* E^{-1} = X_+ \geq 0$ solves the G-DARE (1.1). Note that a $2n \times 2n$ matrix pencil $\mathcal{E} - \lambda\mathcal{F}$ is symplectic

if and only if $\mathcal{E}J\mathcal{E}^T = \mathcal{F}J\mathcal{F}^T$, where $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$. The matrix pencil of the form in (1.9) is symplectic and is said to be a standard symplectic form (SSF), a stronger symplectic property. Being a SSF is the structure we try to preserve in the numerical algorithm (see Chapter 1 for more details of the SSF).

A well-known backward stable approach based on the reordering QZ-algorithm for computing the unique s.p.s.d. solution of DAREs has been proposed by [96]. The associated code, `dare`, has been developed in MATLAB control toolbox [88]. Unfortunately, QZ-like algorithms do not take into account of the symplectic structure, destroying it through the iterative process. Similarly, matrix disk function/inverse free methods [7, 18, 19, 20] have been developed for solving DAREs without preserving the symplectic structure. In Chapter 1, an efficient structure-preserving doubling algorithm (SDA), based on the doubling algorithm [3, 69], has been proposed for solving DAREs, while preserving the SSF at each iterative step. G-DAREs can thus be solved by applying the SDA or other algorithms to the symplectic pencil (1.9). However, the symplectic form in (1.9) requires the explicit inversion of E and R, which may be ill-conditioned.

In this chapter, based on the SDA algorithm in Chapter 1, we develop a generalized structure-preserving doubling algorithm (G-SDA) for solving the G-DARE (1.1). Inversions of ill-conditioned matrices, such as E and R, are circumvented.

2 G-SDA and QR-SWAP Algorithms for G-DAREs

In this section, we develop the G-SDA for solving the G-DARE (1.1). Let

$$\mathcal{M} = \begin{bmatrix} E^{-1}A - E^{-1}BR^{-1}SC^T & 0 \\ -H + CS^TR^{-1}SC^T & I \end{bmatrix}, \quad \mathcal{L} = \begin{bmatrix} I & E^{-1}GE^{-T} \\ 0 & A^TE^{-T} - CS^TR^{-1}B^TE^{-T} \end{bmatrix} \tag{2.1}$$

as given in (1.9), where $H \equiv CQC^T \geq 0$ and $G \equiv BR^{-1}B^T \geq 0$. The pairs (E, A, B) and $(E, A, \widehat{C}_0)$ (with $\widehat{C}_0$ given in (1.4)) are assumed to be (S) and (D), respectively.

Let

$$\widehat{A}_0 \equiv E^{-1}(A - BR^{-1}SC^T), \tag{2.2}$$

$$\widehat{G}_0 \equiv E^{-1}BR^{-1}B^TE^{-T} \rightsquigarrow \widehat{B}_0\widehat{B}_0 \geq 0, \tag{2.3}$$

$$\widehat{H}_0 \equiv CQC^T - CS^TR^{-1}SC^T \rightsquigarrow \widehat{C}_0\widehat{C}_0^T \geq 0. \tag{2.4}$$

Here "$\rightsquigarrow$" denotes the operation which expands a s.p.s.d. matrix into a FRD. The SDA generates the sequences $\{\widehat{A}_k, \widehat{G}_k, \widehat{H}_k\}$:

$$\widehat{A}_{k+1} = \widehat{A}_k(I + \widehat{G}_k\widehat{H}_k)^{-1}\widehat{A}_k, \tag{2.5}$$

$$\widehat{G}_{k+1} = \widehat{G}_k + \widehat{A}_k\widehat{B}_k(I + \widehat{B}_k^T\widehat{C}_k\widehat{C}_k^T\widehat{B}_k)^{-1}\widehat{B}_k^T\widehat{A}_k^T \rightsquigarrow \widehat{B}_{k+1}\widehat{B}_{k+1}^T \geq 0, \tag{2.6}$$

$$\widehat{H}_{k+1} = \widehat{H}_k + \widehat{A}_k^T\widehat{C}_k(I + \widehat{C}_k^T\widehat{B}_k\widehat{B}_k^T\widehat{C}_k)^{-1}\widehat{C}_k^T\widehat{A}_k \rightsquigarrow \widehat{C}_{k+1}\widehat{C}_{k+1}^T \geq 0. \tag{2.7}$$

It was proven in Section 2 of Chapter 1, under conditions (S) and (D), that the sequences $\{\widehat{A}_k, \widehat{G}_k, \widehat{H}_k\}$ converge respectively to zero and the s.p.s.d. solutions of the dual and prime DAREs corresponding to the symplectic pencil (1.9). Consequently, $\{E^{-T}\widehat{H}_kE^{-1}\}$ converges to the s.p.s.d. solution of the G-DARE.

In many applications, the matrices E or R in (2.1) and (2.2)–(2.4) are ill-conditioned, causing numerical instability. We shall modify the SDA (2.5)–(2.7) to the G-SDA for solving G-DAREs. The process utilizes on the basic swapping process [19, Lemma 1]:

Given $E^T \in \mathbb{R}^{r\times r}$, $F^T \in \mathbb{R}^{q\times r}$, let

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} E^T \\ -F^T \end{bmatrix} = \begin{bmatrix} T \\ 0 \end{bmatrix} \tag{2.8}$$

be the QR factorization of $[E^T, -F^T]^T$, where $\bar{F} = Q_{21}^T \in \mathbb{R}^{r\times q}$, $\bar{E} = Q_{22}^T \in \mathbb{R}^{q\times q}$ and T is upper triangular. Then

$$E^{-1}F = \bar{F}\bar{E}^{-1}, \tag{2.9}$$

Here the inverses of E or $\bar{E}$ are merely notational — they do not have to be constructed explicitly unless required by the particular circumstances.

For the G-SDA, we first assume, without loss of generality, that only E is ill-conditioned. We shall later describe a pre-processing step when R is ill-conditioned. Let

$$\begin{aligned} A_0 &\equiv A - BR^{-1}SC^T, \\ G_0 &\equiv BR^{-1}B^T \rightsquigarrow B_0B_0^T \geq 0, \\ H_0 &\equiv CQC^T - CS^TR^{-1}SC^T \rightsquigarrow C_0C_0^T \geq 0. \end{aligned}$$

With

$$G_{k+1} \equiv G_k + A_kE^{-1}B_k(I + B_k^TE^{-T}C_kC_k^TE^{-1}B_k)^{-1}B_k^TE^{-T}A_k^T \rightsquigarrow B_{k+1}B_{k+1}^T,$$

we can rewrite the sequences $\{\widehat{A}_k, \widehat{G}_k, \widehat{H}_k\}$ of (2.5)–(2.7) in the SDA as

$$\widehat{A}_{k+1} \equiv E^{-1}A_{k+1} = E^{-1}A_k(E + B_kB_k^TE^{-T}H_k)^{-1}A_k, \tag{2.10}$$

$$\widehat{G}_{k+1} \equiv E^{-1}G_{k+1}E^{-T} = E^{-1}B_{k+1}B_{k+1}^TE^{-T} \geq 0, \tag{2.11}$$

$$\begin{aligned} \widehat{H}_{k+1} \equiv H_{k+1} &= H_k + A_k^TE^{-T}C_k(I + C_k^TE^{-1}B_kB_k^TE^{-T}C_k)^{-1}C_k^TE^{-1}A_k \\ &\rightsquigarrow C_{k+1}C_{k+1}^T \geq 0. \end{aligned} \tag{2.12}$$

Applying the basic swapping process (2.8)–(2.9) we arrive at the G-SDA which generates the sequences $\{E^{-1}A_k, E^{-1}G_kE^{-T}, H_k\}$ without explicitly inverting E:

$$E^{-1}A_{k+1} = E^{-1}A_k(\bar{E}_k^h)^T(E(\bar{E}_k^h)^T + B_kB_k^T\bar{H}_k)^{-1}A_k, \tag{2.13}$$

$$\begin{aligned} E^{-1}G_{k+1}E^{-T} &= E^{-1}\left[G_k + A_k\bar{B}_k((\bar{E}_k^b)^T\bar{E}_k^b + \bar{B}_k^TC_kC_k^T\bar{B}_k)^{-1}\bar{B}_k^TA_k^T\right]E^{-T} \\ &= E^{-1}B_{k+1}B_{k+1}^TE^{-T} \geq 0, \end{aligned} \tag{2.14}$$

$$\begin{aligned} H_{k+1} &= H_k + A_k^T\bar{C}_k(\bar{E}_k^c(\bar{E}_k^c)^T + \bar{C}_k^TB_kB_k^T\bar{C}_k)^{-1}\bar{C}_k^TA_k \\ &= C_{k+1}C_{k+1}^T \geq 0, \end{aligned} \tag{2.15}$$

where swapping produces, for all $k \geq 0$,

$$E^{-T}H_k = \bar{H}_k(\bar{E}_k^h)^{-T}, \quad E^{-1}B_k = \bar{B}_k(\bar{E}_k^b)^{-1}, \quad E^{-T}C_k = \bar{C}_k(\bar{E}_k^c)^{-T}. \tag{2.16}$$

Notice that other inverses appear in (2.13)–(2.15) and their conditioning will be analyzed in Section 3.

Under conditions (S) and (D), the sequences $\{E^{-1}A_k, E^{-1}G_kE^{-T}, H_k\}$ constructed by (2.13)–(2.16) converge to $\{0, E^{-1}G_*E^{-T}, H_*\}$ in with $E^{-T}H_*E^{-1}$ solving the G-DARE (1.1). By (1.6) the associated optimal control matrix K_s is obtained by

$$\begin{aligned} K_s &= -(R+B^TE^{-T}H_*E^{-1}B)^{-1}(B^TE^{-T}H_*E^{-1}A+SC^T) \\ &= -\left[(\bar{E}_*\bar{E}^TR+\bar{K}_*B)^{-1}\bar{K}_*A+\bar{E}(\bar{E}^T\bar{E}+\bar{B}^TH_*\bar{B})^{-1}\bar{E}^TSC^T\right] \end{aligned}$$

where $B^TE^{-T}=\bar{E}^{-T}\bar{B}^T$ and $(\bar{B}^TH_*)E^{-1}=\bar{E}_*^{-1}\bar{K}_*$ are again computed by swapping.

We now describe the basic G-SDA algorithm for solving G-DARE:

G-SDA Algorithm

Input : E, A, B, C, Q, R, S, τ (a small tolerance);

Note : E is ill-conditioned, R is well-conditioned;

Output : the s.p.s.d. solution X for G-DARE.

Initialize : $A \leftarrow A-BR^{-1}SC^T$, $G \leftarrow BR^{-1}B^T \rightsquigarrow B_0B_0^T \geq 0$,

$H \leftarrow CQC^T - CS^TR^{-1}SC^T \rightsquigarrow C_0C_0^T \geq 0$, $B \leftarrow B_0$, $C \leftarrow C_0$;

Repeat : Swap as in (2.8)–(2.9):

$E^{-T}H = \bar{H}\bar{E}_1^{-T}$, $\quad E^{-1}B = \bar{B}\bar{E}_2^{-1}$, $\quad E^{-T}C = \bar{C}\bar{E}_3^{-T}$;

Compute $W_1 \leftarrow E\bar{E}_1^T + BB^T\bar{H}$, $\quad W_2 \leftarrow \bar{E}_2^T\bar{E}_2 + \bar{B}^TCC^T\bar{B}$,

$W_3 \leftarrow \bar{E}_3\bar{E}_3^T + \bar{C}^TBB^T\bar{C}$;

Solve for $V_1,\ V_2,\ V_3$ from $W_1V_1 = A$, $W_2V_2 = \bar{B}^T$, $W_3V_3 = \bar{C}^T$,

$\widehat{A} \leftarrow A\bar{E}_1^TV_1$, $\widehat{G} \leftarrow G + A\bar{B}V_2A^T$, $\widehat{H} \leftarrow H + A^T\bar{C}V_3A$;

Compute the FRDs: $\widehat{G} \rightsquigarrow \widehat{B}\widehat{B}^T \geq 0$, $\widehat{H} \rightsquigarrow \widehat{C}\widehat{C}^T \geq 0$;

If $\|\widehat{H}-H\| \leq \tau\|\widehat{H}\|$, **Then** $X \leftarrow E^{-T}\widehat{H}E^{-1}$ **Stop**;

Else $A \leftarrow \widehat{A}$, $G \leftarrow \widehat{G}$, $H \leftarrow \widehat{H}$, $B \leftarrow \widehat{B}$, $C \leftarrow \widehat{C}$.

Go to Repeat.

End of G-SDA Algorithm

When R is ill-conditioned, we swap matrices as in (2.8)–(2.9):

$$BR^{-1} = \bar{R}_1^{-1}\bar{B}, \quad R^{-1}(SC^T) = \bar{C}_s^T\bar{R}_2^{-T}. \tag{2.17}$$

Then the pencil $\mathcal{M}_0 - \lambda\mathcal{L}_0$ in (1.8) is equivalent to

$$\begin{bmatrix} \bar{R}_1 A\bar{R}_2^T - \bar{R}_1 B\bar{C}_s^T & 0 \\ -\bar{R}_2(CQC^T)\bar{R}_2^T + \bar{C}_s R\bar{C}_s^T & \bar{R}_2 E^T\bar{R}_1^T \end{bmatrix} - \lambda \begin{bmatrix} \bar{R}_1 E\bar{R}_2^T & \bar{B}R\bar{B}^T \\ 0 & \bar{R}_2 A^T\bar{R}_1^T - \bar{C}_s B^T\bar{R}_1^T \end{bmatrix} \tag{2.18}$$

with the left and right transformations $\mathrm{diag}\{\bar{R}_1, \bar{R}_2\}$ and $\mathrm{diag}\{\bar{R}_2^T, \bar{R}_1^T\}$. Compute the FRDs

$$\bar{B}R\bar{B}^T \rightsquigarrow B_0B_0^T \geq 0\ , \quad \bar{R}_2(CQC^T)\bar{R}_2^T - \bar{C}_s R\bar{C}_s^T \rightsquigarrow C_0C_0^T \geq 0$$

and let

$$A_0 \equiv \bar{R}_1 A\bar{R}_2^T - \bar{R}_1 B\bar{C}_s^T, \quad E_0 \equiv \bar{R}_1(E\bar{R}_2^T) = \bar{R}_1 E_2,$$

the matrix pencil in (2.18) corresponds to a G-DARE with $E = E_0$, $A = A_0$, $B = B_0$, $C = C_0$, $Q = I_n$, $R = I_r$ and $S = 0$. The G-SDA can then be applied, with $R = I_r$ being perfectly conditioned. Let the corresponding s.p.s.d. solution be $E_0^{-T}H_*E_0^{-1}$. Transforming back using (2.17), we see that

$$X_+ = E^{-T}R_2^{-1}H_*R_2^{-T}E^{-1} = E_2^{-T}H_*E_2^{-1} \geq 0$$

solves the original G-DARE (1.1) corresponding to $\mathcal{M}_0 - \lambda\mathcal{L}_0$ in (1.8), and the associated optimal control matrix

$$\begin{aligned} K_s &= -(R + B^TE_2^{-T}H_*E_2^{-T}B)^{-1}(B^TE_2^{-T}H_*E_2^{-1}A + SC^T) \\ &= -\left[(\bar{E}_*\bar{E}_2^TR + \bar{K}_*B)^{-1}\bar{K}_*A + \bar{E}_2(\bar{E}_2^T\bar{E}_2 + \bar{B}_2^TH_*\bar{B}_2)^{-1}\bar{E}_2SC^T\right], \end{aligned}$$

with the help of the swapping $B^TE_2^{-T} = \bar{E}_2^{-T}\bar{B}_2^T$ and $(\bar{B}_2^TH_*)E_2^{-T} = \bar{E}_*^{-1}\bar{K}_*$.

Finally, the s.p.s.d. solution of the G-DARE can be solved by computing the stable deflating subspace of the matrix pencils in (1.7), (1.8) (when R is well-conditioned) or (2.18) by the generalized Schur algorithm. This is equivalent to applying the command `dare` [88] to the G-DARE (1.1). Similarly, the recently developed matrix disk function methods, based on the QR-SWAP process (2.8)–(2.9) [18, 19, 20], can also be applied. For comparison we briefly describe the QR-SWAP matrix disk function method:

QR-SWAP Algorithm [19]

Input : E, A, B, C, Q, R, S, τ (a small tolerance);

Output : the s.p.s.d. solution X for G-DARE;

Initialize : $T \leftarrow 0_n$, $\mathcal{M} \leftarrow \mathcal{M}_0$, $\mathcal{L} \leftarrow \mathcal{L}_0$, where $\mathcal{M}_0$ and $\mathcal{L}_0$ are computed by (1.8);

Repeat : **Compute** the QR-factorization:

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} \mathcal{L} \\ -\mathcal{M} \end{bmatrix} = \begin{bmatrix} \widehat{T} \\ 0 \end{bmatrix};$$

If $\|\widehat{T} - T\| \leq \tau \|\widehat{T}\|$, **Then** solves the least squared problem for X_*:

$$\mathcal{M}(:, 1:n) = \mathcal{M}(:, n+1:2n) X_*;$$

Set $X \leftarrow X_* E^{-1}$,

Else Set $\mathcal{L} \leftarrow Q_{22}\mathcal{L}$, $\mathcal{M} \leftarrow Q_{21}\mathcal{M}$, $T \leftarrow \widehat{T}$;

Go to Repeat.

End of QR-SWAP Algorithm

TSING HUA UNIVERSITY

By (1.6) the associated optimal control matrix is given by

$$K_q = (R + B^T X_* E^{-1} B)^{-1} B^T X_* E^{-1} A = (\bar{E}_* R + \bar{B}_*^T B)^{-1} \bar{B}_*^T A$$

with the help of the swapping $(B^T X_*) E^{-1} = \bar{E}_*^{-1} \bar{B}_*^T$.

3 Conditioning of Inversions in G-SDA

The inversion of the matrices $\left[E(\bar{E}_k^h)^T + B_k B_k^T \bar{H}_k\right]$, $\left[(\bar{E}_k^b)^T \bar{E}_k^b + \bar{B}_k^T C_k C_k^T \bar{B}_k\right]$ and $\left[\bar{E}_k^c (\bar{E}_k^c)^T + \bar{C}_k^T B_k B_k^T \bar{C}_k\right]$ in (2.13)–(2.15) cannot be avoided in the G-SDA. We shall show that the condition numbers of these matrices are small. Let

$$M_k = \begin{bmatrix} E & G_k \\ -H_k & E^T \end{bmatrix} \tag{3.1}$$

where the FRDs $G_k = B_k B_k^T \geq 0$ and $H_k = C_k C_k^T \geq 0$. It is easily checked by using the Sherman-Morrison-Woodbury formula (SMWF) that

$$M_k^{-1} = \begin{bmatrix} (E + G_k E^{-T} H_k)^{-1} & -E^{-1} B_k (I + B_k^T E^{-T} H_k E^{-1} B_k)^{-1} B_k^T E^{-T} \\ E^{-T} C_k (I + C_k^T E^{-1} G_k E^{-T} C_k)^{-1} C_k^T E^{-1} & (E^T + H_k E^{-1} G_k)^{-1} \end{bmatrix}. \tag{3.2}$$

An inspection of (2.10)–(2.12) reveals that the three inverses in (2.13)–(2.15) appear in M_k^{-1}. From the properties of norms, the maximum singular value of M_k^{-1} is a upper bound of the maximum singular value of any of its submatrices. Consequently, we aim to bound $\|M_k^{-1}\|$ from above with a moderate quantity, thus proving that the inversions in (2.13)–(2.15) are well-conditioned.

Since the stabilizability and the detectability of (E, A_k, B_k) and (E, A_k, C_k), respectively, are preserved for all k (see Section 2 of Chapter 1), we shall try to measure the "quality" of stabilizability and detectability and subsequently estimate $\|M_k^{-1}\|$. We shall show that the relevant partial measures of stabilizability and detectability actually improves through the iteration process. Hereafter, $\|\cdot\|$ and $\|\cdot\|_F$ denote the 2-norm and Frobenius norm, respectively. For convenience, we drop the index k in (3.1) and (3.2).

Let (E, A) be a regular pencil and let $\varepsilon > 0$ be a small threshold. Suppose that E is nearly singular and the "large" eigenvalues of (E, A) have only linear divisor. Then there are orthogonal equivalence transformations such that the following equivalence relationship holds:

$$(E, A) \overset{eq.}{\sim} \left(\begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}, \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \right), \quad E_{22}, A_{22} \in \mathbb{R}^{r \times r}, \tag{3.3}$$

in which

$$\max\{\bar{\sigma}(E_{21}), \bar{\sigma}(E_{22})\} \leq \varepsilon \ll \min\{\underline{\sigma}(E_{11}), \underline{\sigma}(A_{22})\}. \tag{3.4}$$

Here $\bar{\sigma}(\cdot) \equiv \|\cdot\|$ and $\underline{\sigma}(\cdot)$ denote, respectively, the maximal and minimal singular values of the given matrix, and r is the number of "large" eigenvalues.

Assume that $r \leq \min\{m, p\}$, where m and p are the rank of the control matrix B and the output matrix C, respectively, as in (1.1), and that $\lambda((E_{11}, A_{11})) \cap \lambda((E_{22}, A_{22})) = \emptyset$

and

$$\frac{\sqrt{r}\varepsilon\|(E_{12}, A_{12})\|_F}{\delta^2} < \frac{1}{4},$$

where $\lambda((\cdot,\cdot))$ denotes the spectrum for the given matrix pair, and

$$\delta = \mathrm{dif}[(E_{11}, A_{11}), (E_{22}, A_{22})] > 0$$

is the difference between the matrix pairs (E_{11}, A_{11}) and (E_{22}, A_{22}) [109, pp. 307]. Then there are P_r, $P_\ell \in \mathbb{R}^{r\times(n-r)}$ with

$$\max\{\|P_r\|, \|P_\ell\|\} \le \frac{2\sqrt{r}\varepsilon}{\delta}$$

satisfying

$$\begin{aligned}
&\begin{bmatrix} I & 0 \\ P_\ell & I \end{bmatrix} \left(\begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}, \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \right) \begin{bmatrix} I & 0 \\ P_r & I \end{bmatrix} \\
= &\left(\begin{bmatrix} \widetilde{E}_{11} & E_{12} \\ 0 & \widetilde{E}_{22} \end{bmatrix}, \begin{bmatrix} \widetilde{A}_{11} & A_{12} \\ 0 & \widetilde{A}_{22} \end{bmatrix} \right),
\end{aligned} \tag{3.5}$$

where

$$\widetilde{E}_{11} \equiv E_{11} + E_{12}P_r, \quad \widetilde{E}_{22} \equiv E_{22} + P_\ell E_{12}, \quad \widetilde{A}_{11} \equiv A_{11} + A_{12}P_r \quad \widetilde{A}_{22} \equiv A_{22} + P_\ell A_{12}. \tag{3.6}$$

It is easily seen that the rows of $[P_\ell, I]$ and the columns of $\begin{bmatrix} -\widetilde{E}_{11}^{-1}E_{12} \\ I \end{bmatrix}$, respectively, form bases of the left and right invariant subspaces of the transformed pencil corresponding to the "large" eigenvalues.

Since (E, A, B) and (E, A, C) are respectively stabilizable and detectable, we have

$$\underline{\sigma}_B \equiv \underline{\sigma}\left(Y_2^T B\right) = \underline{\sigma}\left([P_\ell, I] \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \right) = \underline{\sigma}(\widetilde{B}_2) > 0 \tag{3.7}$$

and

$$\underline{\sigma}_C \equiv \underline{\sigma}\left(C^T X_2\right) = \underline{\sigma}\left([C_1^T, C_2^T] \begin{bmatrix} -\widetilde{E}_{11}^{-1}E_{12} \\ I \end{bmatrix} \right) = \underline{\sigma}(\widetilde{C}_2^T) > 0, \tag{3.8}$$

where $B \equiv \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$, $C^T \equiv \left[C_1^T, C_2^T\right]$, $\widetilde{B}_2 \equiv P_\ell B_1 + B_2$, $\widetilde{C}_2^T \equiv -C_1^T \widetilde{E}_{11} E_{12} + C_2^T$, and X_2 and Y_2 are respectively the unitary bases of the right- and left-eigenvectors corresponding to the large eigenvalues of (E, A). Clearly $\underline{\sigma}_B$ and $\underline{\sigma}_C$ in (3.7) and (3.8), respectively, are partial measures of stabilizability and detectability.

Without loss of generality, we let

$$M = \left[\begin{array}{cc|cc} E_{11} & E_{12} & G_{11} & G_{12} \\ E_{21} & E_{22} & G_{12}^T & G_{22} \\ \hline -H_{11} & -H_{12} & E_{11}^T & E_{21}^T \\ -H_{12}^T & -H_{22} & E_{12}^T & E_{22}^T \end{array}\right], \tag{3.9}$$

where $G_{11} = B_1 B_1^T \geq 0$, $G_{12} = B_1 B_2^T$, $G_{22} = B_2 B_2^T \geq 0$. Multiplying M from the left by

$$L_\ell \equiv \text{diag}\left\{ \begin{bmatrix} I & 0 \\ P_\ell & I \end{bmatrix}, \begin{bmatrix} I & 0 \\ -E_{12}^T \widetilde{E}_{11}^{-T} & I \end{bmatrix} \begin{bmatrix} I & P_r^T \\ 0 & I \end{bmatrix} \right\} \tag{3.10}$$

and from the right by

$$L_r \equiv \text{diag}\left\{ \begin{bmatrix} I & 0 \\ P_r & I \end{bmatrix} \begin{bmatrix} I & -\widetilde{E}_{11}^{-1} E_{12} \\ 0 & I \end{bmatrix}, \begin{bmatrix} I & P_\ell^T \\ 0 & I \end{bmatrix} \right\} \tag{3.11}$$

we obtain, with $\widetilde{E}_{11}$ and $\widetilde{E}_{22}$ given in (3.5),

$$\widetilde{M} = L_\ell M L_r = \left[\begin{array}{cc|cc} \widetilde{E}_{11} & 0 & \widetilde{G}_{11} & \widetilde{G}_{12} \\ 0 & \widetilde{E}_{22} & \widetilde{G}_{12}^T & \widetilde{G}_{22} \\ \hline -\widetilde{H}_{11} & -\widetilde{H}_{12} & \widetilde{E}_{11}^T & 0 \\ -\widetilde{H}_{12}^T & -\widetilde{H}_{22} & 0 & \widetilde{E}_{22}^T \end{array}\right], \tag{3.12}$$

where

$$\begin{bmatrix} \widetilde{G}_{11} & \widetilde{G}_{12} \\ \widetilde{G}_{12}^T & \widetilde{G}_{22} \end{bmatrix} = \begin{bmatrix} B_1 \\ P_\ell B_1 + B_2 \end{bmatrix} \left[B_1^T, B_1^T P_\ell^T + B_2^T\right] = \begin{bmatrix} B_1 \\ \widetilde{B}_2 \end{bmatrix} \left[B_1^T, \widetilde{B}_2^T\right],$$

$$\begin{bmatrix} \widetilde{H}_{11} & \widetilde{H}_{12} \\ \widetilde{H}_{12}^T & \widetilde{H}_{22} \end{bmatrix} = \begin{bmatrix} C_1 \\ -E_{12}^T \widetilde{E}_{11}^{-T} C_1 + C_2 \end{bmatrix} \left[C_1^T, -C_1^T \widetilde{E}_{11}^{-1} E_{12} + C_2^T\right] = \begin{bmatrix} C_1 \\ \widetilde{C}_2 \end{bmatrix} \left[C_1^T, \widetilde{C}_2^T\right].$$

In order to estimate $\|M^{-1}\|$ we first prove the following Lemmas.

Lemma 3.1. *Let E, $G = G^T \geq 0$, $H = H^T \geq 0 \in \mathbb{R}^{n\times n}$.*

(i) *If E is nonsingular, $G \rightsquigarrow BB^T \geq 0$ and $H \rightsquigarrow CC^T \geq 0$, then*

$$\begin{bmatrix} E & G \\ -H & E^T \end{bmatrix}^{-1} = \begin{bmatrix} (E+GE^{-T}H)^{-1} & -E^{-1}G(E^T+HE^{-1}G)^{-1} \\ E^{-T}H(E+GE^{-T}H)^{-1} & (E^T+HE^{-1}G)^{-1} \end{bmatrix}. \tag{3.13}$$

Furthermore, we have

$$\|(E+GE^{-T}H)^{-1}\| \leq \underline{\sigma}(E)^{-1}\left(1+\frac{\|G\|\|H\|}{\underline{\sigma}(E)^2}\right)$$

and

$$\|E^{-1}G(E^T+HE^{-T}G)^{-1}\| \leq \underline{\sigma}(E)^{-2}\|G\|, \quad \|E^{-T}H(E+GE^{-T}H)^{-1}\| \leq \underline{\sigma}(E)^{-2}\|H\|.$$

(ii) *If $G = G^T > 0$ and $H = H^T > 0$, then*

$$\begin{bmatrix} E & G \\ -H & E^T \end{bmatrix}^{-1} = \begin{bmatrix} H^{-1}E^T(G+EH^{-1}E^T)^{-1} & -(H+E^TG^{-1}E)^{-1} \\ (G+EH^{-1}E^T)^{-1} & G^{-1}E(H+E^TG^{-1}E)^{-1} \end{bmatrix}. \tag{3.14}$$

Furthermore, we have

$$\|H^{-1}E^T(G+EH^{-1}E^T)^{-1}\| \leq \frac{\|E\|}{\underline{\sigma}(H)\underline{\sigma}(G)}, \quad \|(H+E^TG^{-1}E)^{-1}\| \leq \frac{1}{\underline{\sigma}(H)}$$

and

$$\|(G+EH^{-1}E^T)^{-1}\| \leq \frac{1}{\underline{\sigma}(G)}, \quad \|G^{-1}E(H+E^TG^{-1}E)^{-1}\| \leq \frac{\|E\|}{\underline{\sigma}(G)\underline{\sigma}(H)}. \tag{3.15}$$

Proof. (i) Equation (3.13) can easily be verified. As in (3.2), using the SMWF twice, we have

$$E^{-1}G(E^T+HE^{-1}G)^{-1} = E^{-1}B(I+B^TE^{-T}HE^{-1}B)^{-1}B^TE^{-T}.$$

Consequently, with $\|(I+B^TE^{-T}HE^{-1}B)^{-1}\| \leq 1$, we have

$$\|E^{-1}G(E^T+HE^{-1}G)^{-1}\| \leq \underline{\sigma}(E)^{-2}\|G\|$$

and similar

$$\|E^{-T}H(E+GE^{-T}H)^{-1}\| \le \underline{\sigma}(E)^{-2}\|H\|.$$

For the diagonal blocks in $\begin{bmatrix} E & G \\ -H & E^T \end{bmatrix}^{-1}$ on the right-hand-side of (3.13), by the SMWF, we have

$$\begin{aligned}(E+GE^{-T}H)^{-1} &= (I+E^{-1}GE^{-T}H)^{-1}E^{-1} \\ &= [I-E^{-1}B(I+B^TE^{-T}HE^{-1}B)^{-1}B^TE^{-T}H]E^{-1}.\end{aligned}$$

Again with $\|(I+B^TE^{-T}HE^{-1}B)^{-1}\| \le 1$, we obtain

$$\|(E+GE^{-T}H)^{-1}\| \le \underline{\sigma}(E)^{-1}\left(1+\frac{\|G\|\|H\|}{\underline{\sigma}(E)^2}\right).$$

(ii) Equation (3.14) can easily be verified. Since $G=G^T>0$ and $H=H^T>0$, we have

$$H^{-1}E^T(G+EH^{-1}E^T)^{-1} = H^{-1}E^TG^{-1/2}(I+G^{-1/2}EH^{-1}E^TG^{-1/2})^{-1}G^{-1/2}$$

and

$$(H+E^TG^{-1}E)^{-1} = H^{-1/2}(I+H^{-1/2}E^TG^{-1}EH^{-1/2})^{-1}H^{-1/2}.$$

Consequently,

$$\|H^{-1}E^T(G+EH^{-1}E^T)^{-1}\| \le \frac{\|E\|}{\underline{\sigma}(H)\underline{\sigma}(G)}, \quad \|(H+E^TG^{-1}E)^{-1}\| \le \frac{1}{\underline{\sigma}(H)}.$$

The inequalities in (3.15) can also be obtained in a similar fashion. □

Lemma 3.2. *Let $\Phi \in \mathbb{R}^{s\times t}$ (w.l.o.g. $s \le t$) with $0 \le \sigma_1 \le \cdots \le \sigma_s \equiv \bar{\sigma}(\Phi)$ singular values. Then it holds*

$$\left\| \begin{bmatrix} I_t & 0 \\ \Phi & I_s \end{bmatrix} \right\| = \left(\frac{2+\sigma_s^2+\sigma_s(4+\sigma_s^2)^{1/2}}{2}\right)^{1/2} \le \bar{\sigma}(\Phi)+1. \tag{3.16}$$

Proof. Applying the singular value decomposition of Φ and the definition

$$\bar{\sigma}(Z) = \lambda_{\max}(Z^TZ)^{1/2},$$

(3.16) follows immediately. □

Theorem 3.3. *Let (E, A) be the matrix pair given by (3.3) satisfying (3.4). Assume that*

$$\lambda((E_{11}, A_{11})) \cap \lambda((E_{22}, A_{22})) = \emptyset, \quad \frac{\sqrt{r}\varepsilon \|(E_{12}, A_{12})\|_F}{\delta^2} < \frac{1}{4}$$

and $P_r, P_\ell \in \mathbb{R}^{r\times(n-r)}$, $r \le \min\{m, p\}$ with

$$\max\{\|P_r\|, \|P_\ell\|\} \le \frac{2\sqrt{r}\varepsilon}{\delta}$$

satisfying (3.5) and (3.6). From (3.7), (3.8) we let

$$e_{11} \equiv \underline{\sigma}(\widetilde{E}_{11}) > 0, \quad g_{22} \equiv \underline{\sigma}_B^2 = \underline{\sigma}(\widetilde{G}_{22}) > 0, \quad h_{22} \equiv \underline{\sigma}_C^2 = \underline{\sigma}(\widetilde{H}_{22}) > 0, \tag{3.17}$$

$$\widetilde{g}_{11} = \bar{\sigma}(\widetilde{G}_{11}), \quad \widetilde{h}_{11} = \bar{\sigma}(\widetilde{H}_{11}), \quad \widetilde{\varepsilon} \equiv \bar{\sigma}(\widetilde{E}_{22}),$$

$$\widetilde{g}_{12} \equiv \bar{\sigma}(\widetilde{G}_{12}), \quad \widetilde{h}_{12} \equiv \bar{\sigma}(\widetilde{H}_{12}), \quad \widetilde{e}_{12} \equiv \bar{\sigma}(E_{12}),$$

$$\eta_g \equiv \widetilde{g}_{11} + \frac{\widetilde{g}_{12}^2}{g_{22}}, \quad \eta_h \equiv \widetilde{h}_{11} + \frac{\widetilde{h}_{12}^2}{h_{22}}, \quad \eta \equiv 1 + \frac{1}{e_{11}^2}\eta_g\eta_h, \tag{3.18}$$

and

$$\kappa \equiv \max\left\{ \left(\left(\eta^2 + \frac{\eta_h^2}{e_{11}^2}\right)\frac{\widetilde{g}_{12}^2}{e_{11}^2 g_{22}^2} + \left(\eta^2 + \frac{\eta_g^2}{e_{11}^2}\right)\frac{\widetilde{h}_{12}^2}{e_{11}^2 h_{22}^2} \right)^{1/2}, \right.$$

$$\left. \frac{1}{g_{22}h_{22}}\left(\frac{\widetilde{h}_{12}^2\widetilde{\varepsilon}^2}{h_{22}^2} + \frac{\widetilde{g}_{12}^2\widetilde{\varepsilon}^2}{g_{22}^2} + \widetilde{g}_{12}^2 + \widetilde{h}_{12}^2 \right)^{1/2} \right\}.$$

If $\kappa\widetilde{\varepsilon} < 1$, then the matrix of the form in (3.9) can be estimated by

$$\|M^{-1}\| \le \max\left\{ \left(\frac{2\eta^2}{e_{11}^2} + \frac{\eta_g^2 + \eta_h^2}{e_{11}^4} \right)^{1/2}, \left(\frac{2\widetilde{\varepsilon}^2 + g_{22}^2 + h_{22}^2}{g_{22}^2 h_{22}^2} \right)^{1/2} \right\} \cdot \frac{\omega}{1 - \kappa\widetilde{\varepsilon}}, \tag{3.19}$$

where

$$\omega = \max\left\{ \left(\frac{2\sqrt{r}\varepsilon}{\delta} + 1 \right), \left(\frac{\widetilde{e}_{12}}{e_{11}} + 1 \right) \right\}^2 \cdot \max\left\{ \left(\frac{\widetilde{g}_{12}}{g_{22}} + 1 \right), \left(\frac{\widetilde{h}_{12}}{h_{22}} + 1 \right) \right\}^2.$$

Proof. Let $\widetilde{M} = L_\ell M L_r$ as in (3.12), where L_ℓ and L_r are given by (3.10) and (3.11), respectively. By assumption (3.17) that $\widetilde{G}_{22}$ and $\widetilde{H}_{22}$ are s.p.d., we eliminate $\widetilde{G}_{12}$ ($\widetilde{G}_{12}^T$) and $\widetilde{H}_{12}$ ($\widetilde{H}_{12}^T$), respectively, by using G_{22} and H_{22} as pivot matrices. Note that G_{22} and H_{22} do not possess full rank when $r > m, p$. Consequently, we have

$$\widetilde{L}_\ell \widetilde{M} \widetilde{L}_r = \widetilde{M}_0 + \widetilde{\mathcal{E}}$$

where

$$\widetilde{M}_0 \equiv \left[\begin{array}{cc|cc} \widetilde{E}_{11} & 0 & \breve{G}_{11} & 0 \\ 0 & \widetilde{E}_{22} & 0 & \widetilde{G}_{22} \\ \hline -\breve{H}_{11} & 0 & \widetilde{E}_{11}^T & 0 \\ 0 & -\widetilde{H}_{22} & 0 & \widetilde{E}_{22}^T \end{array}\right],$$

$$\widetilde{\mathcal{E}} \equiv \operatorname{diag}\left\{\begin{bmatrix} 0 & -\widetilde{G}_{12}\widetilde{G}_{22}^{-1}\widetilde{E}_{22} \\ -\widetilde{H}_{12}^T\widetilde{H}_{22}^{-1}\widetilde{E}_{22} & 0 \end{bmatrix}, \begin{bmatrix} 0 & -\widetilde{E}_{22}^T\widetilde{H}_{22}^{-1}\widetilde{H}_{12}^T \\ -\widetilde{E}_{22}^T\widetilde{G}_{22}^{-1}\widetilde{G}_{12}^T & 0 \end{bmatrix}\right\},$$

$$\widetilde{L}_\ell \equiv \operatorname{diag}\left\{\begin{bmatrix} I & -\widetilde{G}_{12}\widetilde{G}_{22}^{-1} \\ 0 & I \end{bmatrix}, \begin{bmatrix} I & -\widetilde{H}_{12}\widetilde{H}_{22}^{-1} \\ 0 & I \end{bmatrix}\right\}, \tag{3.20}$$

$$\widetilde{L}_r \equiv \operatorname{diag}\left\{\begin{bmatrix} I & 0 \\ -\widetilde{H}_{22}^{-1}\widetilde{H}_{12}^T & I \end{bmatrix}, \begin{bmatrix} I & 0 \\ -\widetilde{G}_{22}^{-1}\widetilde{G}_{12} & I \end{bmatrix}\right\}, \tag{3.21}$$

and $\breve{G}_{11} \equiv \widetilde{G}_{11} - \widetilde{G}_{12}^T\widetilde{G}_{22}^{-1}\widetilde{G}_{12}$, $\breve{H}_{11} \equiv \widetilde{H}_{11} - \widetilde{H}_{12}^T\widetilde{H}_{22}^{-1}\widetilde{H}_{12}$. This implies

$$M^{-1} = L_r\widetilde{L}_r(\widetilde{M}_0 + \widetilde{\mathcal{E}})^{-1}\widetilde{L}_\ell L_\ell = L_r\widetilde{L}_r(I + \widetilde{M}_0^{-1}\widetilde{\mathcal{E}})^{-1}\widetilde{M}_0^{-1}\widetilde{L}_\ell L_\ell. \tag{3.22}$$

Interchanging the second and third row- (and column-) blocks in $\widetilde{M}_0$, we can show the similarity relationship

$$\widetilde{M}_0 \sim \operatorname{diag}\left\{\begin{bmatrix} \widetilde{E}_{11} & \breve{G}_{11} \\ -\breve{H}_{11} & \widetilde{E}_{11}^T \end{bmatrix}, \begin{bmatrix} \widetilde{E}_{22} & \widetilde{G}_{22} \\ -\widetilde{H}_{22} & \widetilde{E}_{22}^T \end{bmatrix}\right\}$$

From Lemma 3.1, we can apply parts (i) and (ii), respectively, to the two diagonal blocks above. One can easily check that $\|\widetilde{M}_0^{-1}\widetilde{\mathcal{E}}\| \le \kappa\widetilde{\varepsilon} < 1$ and obtain

$$\|\widetilde{M}_0^{-1}\| \le \max\left\{\left(\frac{2\eta^2}{e_{11}^2} + \frac{\eta_g^2}{e_{11}^4} + \frac{\eta_h^2}{e_{11}^4}\right)^{1/2}, \left(\frac{2\widetilde{\varepsilon}^2}{g_{22}^2\, h_{22}^2} + \frac{1}{g_{22}^2} + \frac{1}{h_{22}^2}\right)^{1/2}\right\}. \tag{3.23}$$

Furthermore, applying Lemma 3.2 to (3.10), (3.11), (3.20) and (3.21) we have

$$\bar{\sigma}(L_\ell), \bar{\sigma}(L_r) \le \max\left\{\left(\frac{2\sqrt{r}\varepsilon}{\delta} + 1\right), \left(\frac{\widetilde{e}_{12}}{e_{11}} + 1\right)\right\} \tag{3.24}$$

and

$$\bar{\sigma}(\widetilde{L}_\ell), \bar{\sigma}(\widetilde{L}_r) \le \max\left\{\left(\frac{\widetilde{g}_{12}}{g_{22}} + 1\right), \left(\frac{\widetilde{h}_{12}}{h_{22}} + 1\right)\right\}. \tag{3.25}$$

Inequality (3.19) then follows from (3.22)–(3.25). □

Remarks: (1) In Theorem 3.3 we give a upper bound for $\|M^{-1}\|$ in terms of e_{11}, δ and the partial measures of stabilizability and detectability g_{22} and h_{22}. Let ε in (3.4) be chosen so that e_{11} and δ are not too small, and let the systems be reasonably stabilizable and detectable so that the partial measures are reasonably large. In such circumstances, the bound in (3.19) is reasonably moderate for $\|M^{-1}\|$.

(2) The spectrum of (E, A) can spread out continuously in such a way that no small ε together with large e_{11} and δ can be found. In such circumstance, a compromised r has to be chosen so as to optimize the bounds in (3.19).

(3) We can show that the partial measures are not getting worse through the iteration with increasing k. From (2.13)–(2.15), the symmetric G_k and H_k are added to by low-rank updates through the iteration, which increases their ranks as well as their minimum eigenvalues. Recall that the partial measures are really the minimum singular values of $Y_2^T B$ and $C^T X_2$, where Y_2 and X_2 are unitary bases for the eigenvectors corresponding to the large eigenvalues. For various k, they span approximately the null spaces of matrices constructed from E^T and E with components related to E_{21} and E_{22} in (3.3) deleted. Let these null spaces be span respectively by the unitary $\mathcal{N}_r$ and $\mathcal{N}_\ell$. With the help of the properties of norms and variational properties of eigenvalues and singular values, for various values of k, we have

$$\begin{aligned} g_{22} &= Y_2^T B_k B_k^T Y_2 \approx \mathcal{N}_r^T B_k B_k^T \mathcal{N}_r \geq \underline{\sigma}^2(B_k) \geq \underline{\sigma}^2(B_0), \\ h_{22} &= X_2^T C_k C_k^T X_2 \approx \mathcal{N}_\ell^T C_k C_k^T \mathcal{N}_\ell \geq \underline{\sigma}^2(C_k) \geq \underline{\sigma}^2(C_0). \end{aligned}$$

Thus the partial measures for stabilizability and detectability are improving, approximately, through the iteration. The G-SDA should perform well when the original system is reasonably stabilizable and detectable, in the sense that the starting values g_{22} and h_{22} for $k = 0$ are bounded reasonably far away from zero, with a small enough ε and a large enough e_{11}.

(4) The assumption in (3.3) that the large eigenvalues have only linear divisors can be removed. An arbitrarily small perturbation will perturb the matrix pencil (E, A) with defective eigenvalues to one with only linear divisors. The conclusions in the Section then

still hold, effectively via a continuation argument. Alternatively, the system (E, A) can be regularized via feedback [37] before the corresponding G-DARE is solved.

(5) We require that the number of large eigenvalues $r \leq \min\{m, p\}$. This only limits size of ε and the extent to which we can benefit from circumventing the inversion of E. This will not lead to a failure of the G-SDA as E and R, although possibly ill-conditioned, are assumed to be invertible.

4 Numerical Experiments for G-DAREs

For the Tables in the following examples, data for various methods are lists in columns with obvious headings. The heading "`dare`" is for the `dare` command in MATLAB [88] applied to (1.1), and "G-SDA" stands for our G-SDA algorithm. The heading "QR-SWAP" stands for the QR-SWAP algorithm applied to the matrix pencil in (1.7). There is no iteration numbers to report for `dare` and an '$*$' in the Tables indicates a failure to obtain a solution. Besides, we also report the numbers of iterations (no. ite.) for the QR-SWAP and G-SDA algorithms, respectively, and the number of stable closed-loop eigenvalues (no. stab. ev.) in the examples. The symbol $|\lambda^c_{\max}|$ indicates the spectral radius of the closed-loop matrix pair $(E, A + BK)$, i.e.,

$$|\lambda^c_{\max}| = \max\{|\lambda| : \lambda \in \lambda(E, A + BK)\}.$$

We use $\mathrm{trid}(a, b, c)$ to denote the tridiagonal matrix with the main-, sub- and super-diagonal elements being a, b and c, respectively. Also, we denote

$$T_n = \begin{bmatrix} 1 & -1 & -1 & \cdots & -1 \\ 0 & 1 & -1 & \cdots & -1 \\ \vdots & \ddots & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -1 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

For the comparison of residuals computed by these three methods, we use the 'nor-

malized' residual (NRes) formula proposed in [19]

$$\text{NRes} \equiv \frac{\|A^T\tilde{X}A - E^T\tilde{X}E - A^T\tilde{X}B(R+B^T\tilde{X}B)^{-1}B^T\tilde{X}A + H\|}{\|A^T\tilde{X}A\| + \|E^T\tilde{X}E\| + \|A^T\tilde{X}B(R+B^T\tilde{X}B)^{-1}B^T\tilde{X}A\| + \|H\|},$$

where $\tilde{X}$ is an approximate solution.

All computations were performed in MATLAB/version 6.5 on a PC of Intel Pentium-III processor at 866 MHz, with RAM of 768 MB, using IEEE double-precision floating-point arithmetic ($\varepsilon \approx 2.22 \times 10^{-16}$).

Example 1. Consider a linear descriptor system (E, A, B, C) with $n = 6$ and $\text{rank}(B) = \text{rank}(C) = 3$. The system matrices are

$$\begin{aligned}
E &= \text{diag}(1, 10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, 10^{-10}), \quad R = I, \\
A &= \begin{bmatrix} 4.0426 & 3.9258 & 2.6310 & -2.1318 & 5.5853 & -7.1839 \\ 3.5169 & -0.0108 & -1.7188 & -8.5395 & -5.2439 & -0.2965 \\ 4.1518 & 5.7531 & 2.0055 & 4.6018 & 8.2394 & 5.7068 \\ 1.2700 & -7.3705 & -5.6308 & 3.8215 & 8.0503 & 2.2467 \\ 1.5915 & 0.6336 & -2.9188 & 5.2129 & 0.1337 & -6.8345 \\ 4.0271 & -3.9175 & -2.2047 & 2.2661 & 2.8700 & 0.1553 \end{bmatrix}, \\
B^T &= \begin{bmatrix} -0.4820 & -0.4466 & -0.8810 & -0.8007 & 0.4766 & -1.2284 \\ 1.2694 & 0.7538 & -0.8847 & -1.1809 & 0.5286 & 0.3069 \\ -0.6425 & 1.2407 & 0.1126 & 0.7689 & -0.8265 & 0.2993 \end{bmatrix}, \\
C^T &= \begin{bmatrix} 0.3285 & -0.9312 & 1.0424 & 1.1712 & -0.0214 & 0.6355 \\ 0.3685 & 0.6990 & -0.3572 & -0.5304 & -1.7255 & -1.3765 \\ 3.0559 & -2.6376 & -1.2290 & -1.6608 & 0.0370 & 1.3068 \end{bmatrix}.
\end{aligned}$$

In this case, one of the closed-loop eigenvalues achieved by QR-SWAP lies outside the unit circle, with modulus equals 24.255. The numerical results are given in Table 1.

Example 2. In this example, we consider a linear descriptor system (E, A, B, C) with $E = T_n$ and $R = I_n$. System matrices A, B, C are randomly generated with entries of A distributed normally in $[-5, 5]$, and entries of B and C distributed normally in $[-1, 1]$.

	G-SDA	`dare`	QR-SWAP
NRes	1.71×10^{-16}	$*$	1.47×10^{-13}
no. ite.	4	-	5
no. stab. ev.	6	$*$	5
$\lvert\lambda^c_{\max}\rvert$	3.48×10^{-1}	$*$	2.43×10^{1}

Table 1: Results for Example 1.

We set $\text{rank}(B) = \text{rank}(C) = \lceil \frac{n}{2} \rceil$ (the nearest integer $\geq \frac{n}{2}$) for $n = 5, 15, 25, 35, 45$. Note that the matrix E becomes nearly singular for large values of n and its condition number varies from $O(10^1)$ to $O(10^{14})$. For $n = 35$, one of the closed-loop eigenvalues achieved by QR-SWAP lies outside the unit circle, with modulus equals 3.1413. When $n = 45$, four of the closed-loop eigenvalues achieved by QR-SWAP lie outside the unit circle. The numerical results are reported in Table 2.

Example 3. Let $E \in \mathbb{R}^{n\times n}$ be the Frank matrix

$$E = \begin{bmatrix} n & n-1 & n-2 & \cdots & \cdots & 2 & 1 \\ n-1 & n-1 & n-2 & \cdots & \cdots & 2 & 1 \\ 0 & n-2 & n-2 & \cdots & \cdots & 2 & 1 \\ 0 & 0 & n-3 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & & & \ddots & 2 & 1 \\ 0 & 0 & \cdots & & 0 & 1 & 1 \end{bmatrix},$$

and let

$$A = \text{trid}(20, -10, -10) \in \mathbb{R}^{n\times n}, \quad R = I_n.$$

The control matrix B and output matrix C are randomly generated with entries distributed normally in $[-1, 1]$. We set $\text{rank}(B) = \text{rank}(C) = \lceil \frac{n}{2} \rceil$. Note that the matrix E becomes nearly singular for increasing values of n and its condition number varies from $O(1)$ to $O(10^{14})$. The numerical results are reported in Table 3 for $n = 5, 8, 11, 13, 16$.

n	cond(E)		G-SDA	`dare`	QR-SWAP
5	2.9×10^{1}	NRes	9.13×10^{-17}	6.19×10^{-16}	1.70×10^{-14}
		no. ite.	6	-	8
		no. stab. ev.	5	5	5
		$\lvert\lambda^c_{\max}\rvert$	6.32×10^{-1}	6.32×10^{-1}	6.32×10^{-1}
15	9.5×10^{4}	NRes	2.25×10^{-16}	7.90×10^{-16}	8.12×10^{-12}
		no. ite.	5	-	6
		no. stab. ev.	15	15	15
		$\lvert\lambda^c_{\max}\rvert$	1.76×10^{-1}	1.76×10^{-1}	1.76×10^{-1}
25	1.7×10^{8}	NRes	1.04×10^{-16}	$*$	1.09×10^{-12}
		no. ite.	5	-	6
		no. stab. ev.	25	$*$	25
		$\lvert\lambda^c_{\max}\rvert$	1.72×10^{-1}	$*$	1.72×10^{-1}
35	2.4×10^{11}	NRes	2.23×10^{-16}	$*$	5.08×10^{-13}
		no. ite.	5	-	6
		no. stab. ev.	35	$*$	34
		$\lvert\lambda^c_{\max}\rvert$	2.51×10^{-1}	$*$	3.14×10^{0}
45	3.3×10^{14}	NRes	3.11×10^{-16}	$*$	1.18×10^{-13}
		no. ite.	5	-	6
		no. stab. ev.	45	$*$	41
		$\lvert\lambda^c_{\max}\rvert$	5.05×10^{-1}	$*$	2.65×10^{0}

Table 2: Results for Example 2.

For $n = 13, 16$, some closed-loop eigenvalues achieved by QR-SWAP lies outside the unit circle, with moduli up to 2.7023 and 7.6843, respectively.

Example 4. This example is modified from Example 15 in [21], which was presented originally in [96, Example 3]. Here we consider the G-DARE defined by

$$E = \mathrm{diag}(1, 10^{-1}, 10^{-2}, \cdots, 10^{-(n-1)}) \in \mathcal{R}^{n\times n}, \quad A = \mathrm{trid}(0, 0, 1) \in \mathcal{R}^{n\times n},$$

n	cond(E)		G-SDA	`dare`	QR-SWAP
5	6.5×10^2	NRes	4.14×10^{-17}	4.70×10^{-16}	2.71×10^{-14}
		no. ite.	6	-	7
		no. stab. ev.	5	5	5
		$\lvert\lambda^c_{\max}\rvert$	3.62×10^{-1}	3.62×10^{-1}	3.62×10^{-1}
8	2.8×10^5	NRes	3.90×10^{-16}	2.28×10^{-15}	8.70×10^{-10}
		no. ite.	6	-	7
		no. stab. ev.	8	8	8
		$\lvert\lambda^c_{\max}\rvert$	4.79×10^{-1}	4.78×10^{-1}	4.79×10^{-1}
11	3.3×10^8	NRes	9.81×10^{-17}	$*$	5.10×10^{-12}
		no. ite.	7	-	7
		no. stab. ev.	11	$*$	11
		$\lvert\lambda^c_{\max}\rvert$	5.68×10^{-1}	$*$	5.68×10^{-1}
13	5.9×10^{10}	NRes	7.79×10^{-17}	$*$	6.24×10^{-11}
		no. ite.	6	-	7
		no. stab. ev.	13	$*$	12
		$\lvert\lambda^c_{\max}\rvert$	4.46×10^{-1}	$*$	2.70×10^{0}
16	2.3×10^{14}	NRes	2.39×10^{-16}	$*$	2.26×10^{-14}
		no. ite.	6	-	7
		no. stab. ev.	16	$*$	15
		$\lvert\lambda^c_{\max}\rvert$	7.36×10^{-1}	$*$	7.68×10^{0}

Table 3: Results for Example 3.

and

$$B^T = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{1\times n}, \quad R = 1, \quad H = I_n.$$

If $E = \mathrm{diag}(e_{11}, e_{22}, \cdots, e_{nn})$, then it is easily seen that the stabilizing s.p.s.d. solution is

$$X = \mathrm{diag}(x_1, x_2, \cdots, x_n),$$

where $x_1 = 1/e_{11}^2$ and $x_j = (x_{j-1} + 1)/e_{jj}^2$, for $j = 2, \ldots, n$. Note that the closed-loop

eigenvalues are all zero. The numerical results are given in Table 4.

n	cond(E)		G-SDA	`dare`	QR-SWAP
2	10	NRes	1.52×10^{-16}	2.05×10^{-15}	2.95×10^{-17}
		no. ite.	2	-	8
		no. stab. ev.	2	2	2
		$\lvert\lambda^c_{\max}\rvert$	0.00×10^{0}	1.52×10^{-33}	0
4	10^3	NRes	2.32×10^{-16}	2.14×10^{-5}	7.75×10^{-13}
		no. ite.	3	-	7
		no. stab. ev	4	4	4
		$\lvert\lambda^c_{\max}\rvert$	0.00×10^{0}	2.16×10^{-7}	5.51×10^{-7}
6	10^5	NRes	8.15×10^{-17}	$*$	6.58×10^{-1}
		no. ite.	4	-	8
		no. stab. ev.	6	$*$	6
		$\lvert\lambda^c_{\max}\rvert$	0.00×10^{0}	$*$	2.74×10^{-3}
8	10^7	NRes	3.85×10^{-16}	$*$	9.99×10^{-1}
		no. ite.	4	-	11
		no. stab. ev.	8	$*$	6
		$\lvert\lambda^c_{\max}\rvert$	0.00×10^{0}	$*$	1.27×10^{5}
10	10^9	NRes	1.95×10^{-16}	$*$	9.96×10^{-1}
		no. ite.	5	-	13
		no. stab. ev.	10	$*$	6
		$\lvert\lambda^c_{\max}\rvert$	0.00×10^{0}	$*$	5.54×10^{7}

Table 4: Results for Example 4.

Example 5. Here we consider a linear descriptor system (E, A, B, C) with $E = T_n$ and $R = T_m T_m^T$, where $m = \text{rank}(B)$. The matrices A, B, C are randomly generated with entries of A distributed normally in $[-5, 5]$, and entries of B and C distributed normally in $[-1, 1]$. We set $\text{rank}(B) = \text{rank}(C) = \lceil \frac{n}{2} \rceil$ for $n = 5, 15, 25, 35, 45$. Note that the matrices E and R become nearly singular for increasing n and their condition numbers

vary from $O(10^1)$ to $O(10^{15})$. For $n = 35$, one of the closed-loop eigenvalues achieved by QR-SWAP lies outside the unit circle, with modulus equals 20.9885. When $n = 45$, four of the closed-loop eigenvalues achieved by QR-SWAP lie outside the unit circle. The numerical results are reported in Table 5.

n	cond(E)	cond(R)		G-SDA	`dare`	QR-SWAP
5	2.9×10^1	2.9×10^1	NRes	1.97×10^{-16}	3.59×10^{-16}	5.39×10^{-14}
			no. ite.	6	-	8
			no. stab. ev.	5	5	5
			$\lvert\lambda^c_{\max}\rvert$	6.66×10^{-1}	6.66×10^{-1}	6.66×10^{-1}
15	9.5×10^4	1.4×10^5	NRes	7.76×10^{-17}	1.10×10^{-15}	7.29×10^{-12}
			no. ite.	5	-	6
			no. stab. ev.	15	15	15
			$\lvert\lambda^c_{\max}\rvert$	2.40×10^{-1}	2.40×10^{-1}	2.40×10^{-1}
25	1.7×10^8	4.2×10^8	NRes	4.84×10^{-16}	$*$	2.74×10^{-12}
			no. ite.	6	-	6
			no. stab. ev.	25	$*$	25
			$\lvert\lambda^c_{\max}\rvert$	2.21×10^{-1}	$*$	2.21×10^{-1}
35	2.4×10^{11}	8.6×10^{11}	NRes	2.25×10^{-16}	$*$	2.95×10^{-12}
			no. ite.	6	-	6
			no. stab. ev.	35	$*$	34
			$\lvert\lambda^c_{\max}\rvert$	3.12×10^{-1}	$*$	2.10×10^1
45	3.3×10^{14}	1.5×10^{15}	NRes	4.96×10^{-16}	$*$	4.35×10^{-14}
			no. ite.	6	-	6
			no. stab. ev.	45	$*$	41
			$\lvert\lambda^c_{\max}\rvert$	8.60×10^{-1}	$*$	4.11×10^1

Table 5: Results for Example 5.

5 Conclusions

We have develop the G-SDA algorithm which solves G-DAREs with ill-conditioned R and E. Inversions of ill-conditioned matrices are circumvented via well selected swapping of matrix products and applications of the SMWF. Numerical results show that the G-SDA is competitive with QR-SWAP and `dare`, out-performing the other algorithms in the selected set of test examples. The advantage of our structure-preserving algorithm is evident from the absence of unstable closed-loop eigenvalues at the end of the iterative process, contrasting results from QR-SWAP. The corresponding solution from QR-SWAP may be accurate in the sense of a small residual, it is obviously useless in the sense of stabilizing the closed-loop system. The MATLAB command `dare` failed frequently for many ill-conditioned examples. Apart from having superior accuracy, convergence and structure-preserving properties, the operation count (per iteration) for G-SDA is a small fraction of those for the other algorithms, analogous to the superiority of the SDA for DAREs proposed in Chapter 1. This efficiency is the consequence of the fact that the G-SDA operates in $\mathbb{R}^{n\times n}$ while QR-SWAP and `dare` works with matrices of higher dimensions.

TSING HUA UNIVERSITY

Chapter 4

Balanced Realization of Periodic Descriptor Systems

1 Introduction

In the second-half of the last century, the development of systems and control theory, together with the achievements of digital control and signal processing, has set the stage for renewed interests in the study of periodic systems, both in continuous and discrete time; see, e.g., [87, 130, 118, 47, 57, 52] and the survey papers [28, 29]. This has been amplified by specific application demands in the aerospace realm [68, 89, 67], computer control of industrial processes [30] and communication systems [117, 47, 116, 129]. The number of contributions on linear time-varying discrete-time periodic systems has been increasing in recent times; see, e.g., [53, 62, 72, 121, 123, 125] and the references therein. This increasing interest in periodic systems has also been motivated by the large variety of processes that can be modelled through linear discrete-time periodic systems (e.g., multirate sampled-data systems, chemical processes, periodically time-varying filters and networks, and seasonal phenomena [26, 28, 31, 55, 83, 101, 128]).

We consider here periodic time-varying descriptor systems of the form

$$E_k x_{k+1} = A_k x_k + B_k u_k, \quad y_k = C_k x_k, \quad k \in \mathbb{Z}, \tag{1.1}$$

where the matrices E_k, $A_k \in \mathbb{R}^{n\times n}$, $B_k \in \mathbb{R}^{n\times m}$, $C_k \in \mathbb{R}^{p\times n}$ are periodic with period $K \geq 1$, i.e., $E_k = E_{k+K}$, $A_k = A_{k+K}$, $B_k = B_{k+K}$, $C_k = C_{k+K}$, and the matrices E_k are allowed to be singular for all k. Recently, this class of periodic descriptor systems (1.1) is discussed and studied extensively in the problem of solvability and conditionability [107], the computation of H_∞-norm and system zeros [106, 127], and the compensating and regularization problems for periodic descriptor systems [35, 73].

It is well known that the dynamics of the discrete-time periodic descriptor system (1.1) depend critically on the regularity and the eigenstructure of the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ which satisfy the homogeneous systems of (1.1):

$$E_k x_{k+1} = A_k x_k, \quad k \in \mathbb{Z}. \tag{1.2}$$

The matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ are said to be regular when $\det[C((\alpha_k, \beta_k)_{k=0}^{K-1})] \not\equiv 0$, where

$$C((\alpha_k, \beta_k)_{k=0}^{K-1}) \equiv \begin{bmatrix} \alpha_0 E_0 & 0 & \cdots & 0 & -\beta_0 A_0 \\ -\beta_1 A_1 & \alpha_1 E_1 & & & 0 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & 0 \\ 0 & & 0 & -\beta_{K-1} A_{K-1} & \alpha_{K-1} E_{K-1} \end{bmatrix}, \tag{1.3}$$

in which α_k, β_k are complex variables for $k = 0, \ldots, K-1$.

Definition 1.1. *[80] Let $\{(E_k, A_k)\}_{k=0}^{K-1}$ be $n \times n$ regular matrix pairs. If there exist complex numbers $\alpha_0, \cdots, \alpha_{K-1}$, $\beta_0, \cdots, \beta_{K-1}$ which satisfy*

$$\det[C((\alpha_k, \beta_k)_{k=0}^{K-1})] = 0, \quad \left(\prod_{k=0}^{K-1} \alpha_k, \prod_{k=0}^{K-1} \beta_k \right) \equiv (\pi_\alpha, \pi_\beta) \neq (0, 0) \tag{1.4}$$

then (π_α, π_β) is an eigenvalue pair of $\{(E_k, A_k)\}_{k=0}^{K-1}$.

Note that if (π_α, π_β) is an eigenvalue pair of $\{(E_k, A_k)\}_{k=0}^{K-1}$, then (π_α, π_β) and $(\tau\pi_\alpha, \tau\pi_\beta)$ represent the same eigenvalue for any nonzero τ. If $\pi_\beta \neq 0$, then $\lambda = \pi_\alpha / \pi_\beta$ is a finite eigenvalue; otherwise $(\pi_\alpha, 0)$ represents an infinite eigenvalue. The spectrum, or the set of all eigenvalue pairs, of $\{(E_k, A_k)\}_{k=0}^{K-1}$ is denoted by $\sigma\left(\{(E_k, A_k)\}_{k=0}^{K-1}\right)$. We shall assume throughout this chapter that the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ are regular, and also use the notation $\sigma(M)$ to denote the spectrum of a square matrix M.

It is easily seen that the determinant of $C((\alpha_k, \beta_k)_{k=0}^{K-1})$ is a homogeneous polynomial in π_α and π_β of degree n of the form

$$\sum_{k=0}^{n} c_k \pi_\alpha^k \pi_\beta^{n-k}, \tag{1.5}$$

where $c_0, \cdots, c_n$ are complex numbers uniquely determined by $\{(E_k, A_k)\}_{k=0}^{K-1}$. For the regular matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$, at least one of the c_k's is nonzero, and hence we see from Definition 1.1 that there are exact n eigenvalue pairs (counting multiplicity) for $\{(E_k, A_k)\}_{k=0}^{K-1}$.

It was shown in [107] that the solvability of (1.2) is equivalent to the condition that the pencil

$$\alpha\mathcal{E} - \beta\mathcal{A} := \begin{bmatrix} \alpha E_0 & 0 & \cdots & 0 & -\beta A_0 \\ -\beta A_1 & \alpha E_1 & & & 0 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & 0 \\ 0 & & 0 & -\beta A_{K-1} & \alpha E_{K-1} \end{bmatrix} \tag{1.6}$$

is regular i.e. $\det(\alpha\mathcal{E} - \beta\mathcal{A}) \not\equiv 0$. From (1.5) it is easy to check that

$$\sigma\left(\{(E_k, A_k)\}_{k=0}^{K-1}\right) = \left\{(\alpha^K, \beta^K) \mid \det(\alpha\mathcal{E} - \beta\mathcal{A}) = 0\right\}. \tag{1.7}$$

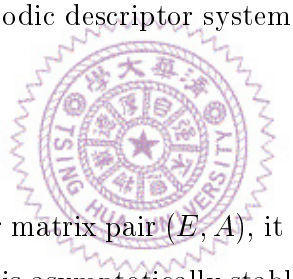
Hence, from (1.7), the solvability of (1.2) is equivalent to the regularity of $\{(E_k, A_k)\}_{k=0}^{K-1}$.

For discrete-time descriptor systems, the concepts of reachability and observability Gramians, causal and noncausal Hankel singular values, and balanced realization are well-established [17, 114]. Moreover, numerical methods are proposed in [110] to solve the projected generalized Lyapunov equations for continuous-time descriptor systems. However, to our best knowledge, similar results have not been developed for periodic descriptor systems.

In summary, there are three main contributions from this chapter. First, in Section 3, we give a set of necessary and sufficient conditions of complete reachability and observability for the periodic time-varying descriptor system (1.1). Second, with the aid of the fundamental matrices $\varphi_{i,j}$ defined as in (2.6), the reachability/observability Gramians and their corresponding projected generalized discrete-time periodic Lyapunov equations (GDPLE) are derived in terms of the original system matrices E_k, A_k, B_k and C_k, $k = 0, 1, \ldots, K-1$, respectively. These fundamental matrices play an important role here and are not natural extension of those defined for the descriptor system with

period $K = 1$ [110, 114]. Third, in Sections 6 and 7, Hankel singular values and balanced realization are discussed, for the first time, for completely reachable and observable periodic descriptor systems. These concepts are likely to be crucial in the model reduction problem of periodic descriptor systems.

This chapter is organized as follows. Section 2 contains some notations and definitions, as well as some preliminary results. In Section 3 the necessary and sufficient conditions are derived for complete reachability and observability of periodic descriptor systems, respectively. With these equivalent conditions, the periodic reachability and observability Gramians, which satisfy some generalized periodic Lyapunov equations, are developed in Section 4. In Section 5 we propose a numerical method for solving these equations under the assumption of pd-stability. A numerical example is given to illustrate its feasibility and reliability. The concept of Hankel singular values is generalized for periodic descriptor systems in Section 6. The problem of balanced realization for the completely reachable and completely observable periodic descriptor systems is discussed in Section 7.



2 Preliminaries

For period $K = 1$ and a regular matrix pair (E, A), it is well known that the discrete-time descriptor system (E, A, B, C) is asymptotically stable if and only if all finite eigenvalues of (E, A) lie inside the unit circle [48, 111, 112]. Similarly, the asymptotic stability of the periodic descriptor system (1.1) can be characterized in terms of the spectrum of the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$.

Definition 2.1. *Let $\{(E_k, A_k)\}_{k=0}^{K-1}$ be $n \times n$ regular matrix pairs. The periodic descriptor system (1.1) is asymptotically stable if and only if all finite eigenvalues of the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ lie inside unit circle. The periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ are called pd-stable if the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ are regular and all their finite eigenvalues lie inside the unit circle.*

In a similar fashion to the Kronecker canonical form for a regular matrix pair, we can transform regular periodic matrix pairs into periodic Kronecker canonical forms [73].

Lemma 2.1. *Suppose that the periodic matrix pairs* $\{(E_k, A_k)\}_{k=0}^{K-1}$ *in systems (1.1) are regular. Then for* $k = 0, \ldots, K-1$, *there exist nonsingular matrices* X_k *and* Y_k *such that*

$$X_k E_k Y_{k+1} = \begin{bmatrix} I & 0 \\ 0 & E_k^b \end{bmatrix}, \quad X_k A_k Y_k = \begin{bmatrix} A_k^f & 0 \\ 0 & I \end{bmatrix}, \tag{2.1}$$

where $Y_K \equiv Y_0$, $A_{k+K-1}^f A_{k+K-2}^f \cdots A_k^f \equiv J_k$ *is an* $n_1 \times n_1$ *Jordan matrix corresponding to the finite eigenvalues,* $E_k^b E_{k+1}^b \cdots E_{k+K-1}^b \equiv N_k$ *is an* $n_2 \times n_2$ *nilpotent Jordan matrix corresponding to the infinite eigenvalues, and* $n = n_1 + n_2$.

Remark. If ν_k is the nilpotency of the nilpotent matrix N_k for $k = 0, 1, \ldots, K-1$, then these K values are defined as the indices [73] of regular periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$. Hence we define the index of the periodic descriptor system (1.1) as $\nu \equiv \max\{\nu_0, \nu_1, \cdots, \nu_{K-1}\}$. We say that the periodic descriptor system (1.1) is of index at most 1 if $\nu \leq 1$, i.e., E_k are all nonsingular or $N_k = 0$ for all k.

For each $k \in \mathbb{Z}$, we let

$$x_k = Y_k \begin{bmatrix} x_k^f \\ x_k^b \end{bmatrix} \begin{matrix} \}n_1 \\ \}n_2 \end{matrix}, \quad X_k B_k = \begin{bmatrix} B_k^f \\ B_k^b \end{bmatrix} \begin{matrix} \}n_1 \\ \}n_2 \end{matrix}, \quad C_k Y_k = \begin{bmatrix} \underset{n_1}{C_k^f} & \underset{n_2}{C_k^b} \end{bmatrix}, \tag{2.2}$$

and by using Lemma 2.1 we can decompose the original system (1.1) into forward and backward periodic subsystems, respectively:

$$x_{k+1}^f = A_k^f x_k^f + B_k^f u_k, \quad y_k^f = C_k^f x_k^f, \tag{2.3}$$

$$E_k^b x_{k+1}^b = x_k^b + B_k^b u_k, \quad y_k^b = C_k^b x_k^b, \tag{2.4}$$

with $y_k = y_k^f + y_k^b, \quad k \in \mathbb{Z}$.

Notice that the state transition matrix of the forward subsystem (2.3) equals $\Phi_f(i, j) = A_{i-1}^f A_{i-1}^f \cdots A_j^f$ when $i > j$ with $\Phi_f(i, i) := I_{n_1}$. The state transition matrix of the backward subsystem (2.4) is $\Phi_b(i, j) = E_i^b E_{i+1}^b \cdots E_{j-1}^b$ when $i < j$ with $\Phi_b(i, i) := I_{n_2}$. The state transition matrix over one period $\Phi_f(\tau+K, \tau) \in \mathbb{R}^{n_1 \times n_1}$ is called the monodromy matrix of the forward subsystem (2.3) at time τ. It is well known that its eigenvalues, called the characteristic multipliers, are independent of τ [122, 81].

For $k = 0, 1, \ldots, K-1$, the $n \times n$ matrices

$$P_r(k) = Y_k \begin{bmatrix} I_{n_1} & 0 \\ 0 & 0 \end{bmatrix} Y_k^{-1}, \quad P_l(k) = X_k^{-1} \begin{bmatrix} I_{n_1} & 0 \\ 0 & 0 \end{bmatrix} X_k, \tag{2.5}$$

are respectively the spectral projections onto the kth right and left deflating subspaces of the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ corresponding to the finite eigenvalues. Moreover, the fundamental matrices $\varphi_{i,j}$ $(i, j \in \mathbb{Z})$ of the periodic descriptor system (1.1) are defined by

$$\varphi_{i,j} = \begin{cases} Y_i \begin{bmatrix} \Phi_f(i, j+1) & 0 \\ 0 & 0 \end{bmatrix} X_j, & \text{if } i > j, \\ Y_i \begin{bmatrix} 0 & 0 \\ 0 & -\Phi_b(i, j) \end{bmatrix} X_j, & \text{if } i \leq j. \end{cases} \tag{2.6}$$

These matrices play an essential role for the periodic discrete-time descriptor system (1.1). For the discrete-time descriptor system with period $K = 1$, these fundamental matrices coincide with the coefficient matrices of the Laurent expansion of the generalized resolvent $(\lambda E - A)^{-1}$ at infinity [79, 114].

3 Complete Reachability and Observability

In this Section we shall give a characterization of complete reachability and observability for the periodic discrete-time descriptor systems (1.1).

Definition 3.1. (i) *The periodic descriptor system (1.1) is reachable at time t if for any state $\bar{x} \in \mathbb{R}^n$, there exist two integers s, ℓ with $s < t < \ell$ and a set of control inputs $\{u_i\}_{i=s}^{\ell}$ which carry $x_s = 0$ into $x_t = \bar{x}$. The periodic descriptor system (1.1) is called completely reachable if it is reachable at all time t.*

(ii) *The forward subsystem (2.3) is reachable at time t if for any state $\bar{\xi}_1 \in \mathbb{R}^{n_1}$, there exists an integer s with $s < t$ and a set of control inputs $\{u_i\}_{i=s}^{t-1}$ which carry $x_s^f = 0$ into $x_t^f = \bar{\xi}_1$. The periodic subsystem (2.3) is called completely reachable if it is reachable at all time t.*

(iii) *The backward subsystem (2.4) is reachable at time t if for any state $\bar{\xi}_2 \in \mathbb{R}^{n_2}$, there exists an integer ℓ with $\ell > t$ and a set of control inputs $\{u_i\}_{i=t}^{\ell}$ such that $x_t^b = \bar{\xi}_2$. The periodic subsystem (2.4) is completely reachable if it is reachable at all time t.*

Remark. It is easily seen from Definition 3.1 that the periodic discrete-time descriptor system (1.1) is completely reachable if and only if both its forward and backward subsystems are completely reachable.

Theorem 3.1 (Forward Reachability). *The following statements are equivalent.*

(a) *The forward subsystem (2.3) is completely reachable.*

(b) *For $t = 0, 1, 2, \cdots, K-1$, the matrices*

$$\mathcal{R}^f(t) \equiv \left[B_{t-1}^f,\ A_{t-1}^f B_{t-2}^f,\ \cdots,\ \Phi_f(t, t-n_1K+1)B_{t-n_1K}^f\right]$$

have full row rank .

(c) *For $t = 0, 1, 2, \cdots, K-1$, and*

$$\mathcal{B}_t^f \equiv \left[B_{t-1}^f,\ A_{t-1}^f B_{t-2}^f,\ A_{t-1}^f A_{t-2}^f B_{t-3}^f,\ \cdots,\ \Phi_f(t, t-K+1)B_{t-K}^f\right]\ ,$$

the matrices

$$\left[\mathcal{B}_t^f,\ \Phi_f(t, t-K)\mathcal{B}_t^f,\ (\Phi_f(t, t-K))^2\mathcal{B}_t^f,\ \cdots,\ (\Phi_f(t, t-K))^{n_1-1}\mathcal{B}_t^f\right]$$

have full row rank.

(d) *For $\prod_{i=0}^{K-1} \alpha_i \in \sigma(\Phi_f(K, 0))$, the matrix*

$$U^f(\alpha_0, \cdots, \alpha_{K-1}) \equiv \left[\begin{array}{ccccc|ccccc} \alpha_0 I & 0 & \cdots & 0 & -A_0^f & B_0^f & & & & \\ -A_1^f & \alpha_1 I & \ddots & & 0 & & B_1^f & & & \\ 0 & -A_2^f & \ddots & \ddots & \vdots & & & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & 0 & & & & \ddots & \\ 0 & \cdots & 0 & -A_{K-1}^f & \alpha_{K-1} I & & & & & B_{K-1}^f \end{array}\right]$$

has full row rank.

(e) *For* $t = 0, 1, 2, \cdots, K-1$,

$$y^T \Phi_f(t+K, t) = \lambda y^T \ \textit{and}\ y^T \Phi_f(t, j) B^f_{j-1} = 0 \ \textit{for}\ j = t-K+1, \cdots, t-1, t$$

imply $y = 0$.

Proof. (a) $\Rightarrow$ (e): Suppose the statement (a) is true. For any $t \in \{0, 1, \cdots, K-1\}$, assume that

$$y^T \Phi_f(t+K, t) = \lambda y^T \ \text{and}\ y^T \Phi_f(t, j) B^f_{j-1} = 0 \ \text{for}\ j = t-K+1, \cdots, t-1, t. \quad (3.1)$$

Since the forward subsystem (2.3) is reachable at time t, there exist an integer s with $s < t$ and control inputs u_i, $s \le i \le t-1$, which carry $x^f_s = 0$ into $x^f_t = y$. Thus, we have

$$y = x^f_t = \sum_{i=s}^{t-1} \Phi_f(t, i+1) B^f_i u_i.$$

Moreover, from the assumptions (3.1), it follows that

$$y^T y = y^T \sum_{i=s}^{t-1} \Phi_f(t, i+1) B^f_i u_i = 0.$$

Therefore, $y = 0$ and hence the condition (e) holds.

(e) $\Rightarrow$ (d): Assume that the condition (e) holds, and let vectors $y_0, y_1, \cdots, y_{K-1} \in \mathbb{R}^{n_1}$ satisfy

$$(y_0^T, y_1^T, \cdots, y_{K-1}^T) U^f(\alpha_0, \alpha_1, \cdots, \alpha_{K-1}) = 0,$$

or

$$\begin{cases} \alpha_0 y_0^T = y_1^T A^f_1 \\ \alpha_1 y_1^T = y_2^T A^f_2 \\ \cdots\cdots\cdots\cdots \\ \alpha_{K-2} y_{K-2}^T = y_{K-1}^T A^f_{K-1} \\ \alpha_{K-1} y_{K-1}^T = y_0^T A^f_0 \end{cases} \quad \text{and} \quad \begin{cases} y_0^T B^f_0 = 0 \\ y_1^T B^f_1 = 0 \\ \cdots\cdots\cdots \\ y_{K-2}^T B^f_{K-2} = 0 \\ y_{K-1}^T B^f_{K-1} = 0. \end{cases} \quad (3.2)$$

Notice that for the vector y_{K-1}, it can be easily checked from (3.2), for $j = -K+1, \cdots, -2, -1, 0$, that

$$\begin{aligned} y_{K-1}^T \Phi_f(K,0) &= y_{K-1}^T A_{K-1}^f A_{K-2}^f \cdots A_0^f \\ &= (\alpha_{K-2}\alpha_{K-3}\cdots\alpha_0\alpha_{K-1}) y_{K-1}^T, \end{aligned}$$

and

$$y_{K-1}^T \Phi_f(0,j) B_{j-1}^f = 0 \ .$$

By condition (e), if the product $\alpha_{K-2}\alpha_{K-3}\cdots\alpha_0\alpha_{K-1} \in \sigma(\{(E_k, A_k)\}_{k=0}^{K-1})$, we then have $y_{K-1} = 0$. Similarly, it can be shown that $y_{K-2} = y_{K-3} = \cdots = y_0 = 0$. Therefore, the matrix $U^f(\alpha_0, \cdots, \alpha_{K-1})$ has full row rank and condition (d) is proved.

(d) $\Rightarrow$ (c): Suppose that (d) holds. It suffices to prove condition (c) for time instant $t = 0$. Since condition (d) holds, it follows that

$$U^f(1, \cdots, 1, \lambda) = \left[\begin{array}{ccccc|cccc} I & 0 & \cdots & 0 & -A_0^f & B_0^f & & & \\ -A_1^f & I & \ddots & & 0 & & B_1^f & & \\ 0 & -A_2^f & \ddots & \ddots & \vdots & & & \ddots & \\ \vdots & \ddots & \ddots & \ddots & 0 & & & & \ddots \\ 0 & \cdots & 0 & -A_{K-1}^f & \lambda I & & & & B_{K-1}^f \end{array}\right] \tag{3.3}$$

has full row rank for all $\lambda \in \sigma(\Phi_f(K,0))$. By elementary row operations, the matrix $U^f(1, \cdots, 1, \lambda)$ can be transformed to

$$\tilde{U}^f \equiv \left[\begin{array}{ccccc|cccccc} I & * & \cdots & \cdots & * & * & & & & & \\ 0 & I & \ddots & & \vdots & \vdots & \ddots & & & & \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & & \ddots & & & \\ \vdots & & \ddots & I & * & * & \cdots & \cdots & * & & \\ 0 & \cdots & \cdots & 0 & \lambda I - \Phi_f(K,0) & \Phi_f(K,1)B_0^f & \cdots & \cdots & A_{K-1}^f B_{K-2}^f & B_{K-1}^f \end{array}\right],$$

which is of full row rank for all $\lambda \in \sigma(\Phi_f(K,0))$. Equivalently, the last row blocks of $\tilde{U}^f$ has full row rank for all $\lambda \in \sigma(\Phi_f(K,0))$, i.e.,

$$\text{rank} \left[\begin{array}{c|c} \lambda I - \Phi_f(K,0) & \mathcal{B}_0^f \end{array}\right] = n_1 \quad \text{for } \lambda \in \sigma(\Phi_f(K,0)),$$

where the matrix $\mathcal{B}_0^f$ is defined in the (c). This proves condition (c) for $t = 0$. Similar arguments apply for $1 \le t \le K-1$.

(c) $\Rightarrow$ (b): This follows from the periodicity of the matrices A_k^f and B_k^f, i.e., $A_{k+K}^f = A_k^f$ and $B_{k+K}^f = B_k^f$ for all k.

(b) $\Rightarrow$ (a): Assume that condition (b) holds. For any time $t \in \mathbb{Z} \pmod K$ and any given state $\bar{\xi}_1 \in \mathbb{R}^{n_1}$, there exist an integer $s \equiv t - n_1 K$ and a set of control inputs u_i, $s \le i \le t-1$, which satisfy

$$\sum_{i=s}^{t-1} \Phi_f(t, i+1) B_i^f u_i = \bar{\xi}_1,$$

since $\mathcal{R}^f(t)$ has full row rank. With these control inputs, the given state $\bar{\xi}_1$ can be reached at time t from the initial state $x_s^f = 0$, and hence the complete reachability of the forward subsystem (2.3) is proved. □

Theorem 3.2 (Backward Reachability). *The following statements are equivalent.*

(a) *The backward subsystem (2.4) is completely reachable.*

(b) *For $t = 0, 1, 2, \cdots, K-1$, the matrices*

$$\mathcal{R}^b(t) \equiv \left[B_t^b,\ E_t^b B_{t+1}^b,\ \cdots,\ \Phi_b(t, t+\nu K-1) B_{t+\nu K-1}^b \right]$$

have full row rank.

(c) *For $t = 0, 1, 2, \cdots, K-1$, and*

$$\mathcal{B}_t^b \equiv \left[B_t^b,\ E_t^b B_{t+1}^b,\ \cdots,\ E_t^b E_{t+1}^b \cdots E_{t+K-2}^b B_{t+K-1}^b \right],$$

the matrices $\left[N_t,\ \mathcal{B}_t^b \right]$ have full row rank.

(d) *The pair $(\mathcal{E}_b, \mathcal{B}_b)$ is reachable, where*

$$\mathcal{E}_b \equiv \begin{bmatrix} 0 & E_0^b & & & \\ 0 & 0 & E_1^b & & \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & 0 & & \ddots & E_{K-2}^b \\ E_{K-1}^b & 0 & \cdots & \cdots & 0 \end{bmatrix} \quad \text{and} \quad \mathcal{B}_b \equiv \begin{bmatrix} B_0^b & & & & \\ & B_1^b & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & B_{K-1}^b \end{bmatrix}. \tag{3.4}$$

Proof. (a) $\Leftrightarrow$ (b): For any time $t \in \mathbb{Z}$, it can be easily seen that

$$x_t^b = -\sum_{i=t}^{t+\nu K-1} \Phi_b(t,i) B_i^b u_i = -\mathcal{R}^b(t) \begin{bmatrix} u_t \\ u_{t+1} \\ \vdots \\ u_{t+\nu K-1} \end{bmatrix}.$$

Therefore, any given state $\bar{\xi}_2 \in \mathbb{R}^{n_2}$ can be reached at time t, i.e., $x_t^b = \bar{\xi}_2$, through a set of control inputs $\{u_i\}_{i=t}^{t+\nu K-1}$ if and only if the matrix $\mathcal{R}^b(t)$ is of full row rank.

(b) $\Leftrightarrow$ (c): Since $N_t = \Phi_b(t, t+K) = E_t^b E_{t+1}^b \cdots E_{t+K-1}^b$ and $N_t^\nu = 0$ for $t = 0, 1, \ldots, K-1$, it follows that

$$\mathcal{R}^b(t) = \left[\mathcal{B}_t^b,\ N_t\mathcal{B}_t^b,\ \cdots,\ N_t^{\nu-1}\mathcal{B}_t^b\right].$$

Thus, $\operatorname{rank}(\mathcal{R}^b(t)) = n_2$ if and only if $\operatorname{rank}\left[\lambda I - N_t,\ \mathcal{B}_t^b\right] = n_2$ for any $\lambda \in \sigma(N_t)$. Since the matrix N_t is nilpotent, $\sigma(N_t) = \{0\}$, and hence $\operatorname{rank}(\mathcal{R}^b(t)) = n_2$ if and only if $\operatorname{rank}\left[-N_t,\ \mathcal{B}_t^b\right] = n_2$. Equivalently, $\operatorname{rank}\left[N_t,\ \mathcal{B}_t^b\right] = n_2$.

(b) $\Leftrightarrow$ (d): Notice that the matrix $\mathcal{E}_b$ is nilpotent with the property that $\mathcal{E}_b^\nu = 0$ and $\mathcal{E}_b^{\nu-1} \neq 0$. It is well known that the pair $(\mathcal{E}_b, \mathcal{B}_b)$ is reachable if and only if $\mathbb{B}_b \equiv \left[\mathcal{B}_b,\ \mathcal{E}_b\mathcal{B}_b,\ \cdots,\ \mathcal{E}_b^{\nu-1}\mathcal{B}_b\right]$ has full row rank. Furthermore, it can be checked that the row blocks of the matrix $\mathbb{B}_b$ are just $\mathcal{R}^b(t)$ with different t. Therefore, statements (b) and (d) are equivalent. □

Definition 3.2. (i) *The periodic descriptor system (1.1) is observable at time t if there exist two integers s, ℓ with $s < t < \ell$ such that any state at time t can be determined from knowledge of $\{y_i\}_{i=s}^{\ell}$ and $\{u_i\}_{i=s}^{\ell}$. The periodic descriptor system (1.1) is called completely observable if it is observable at all time t.*

(ii) *The forward subsystem (2.3) is observable at time t if there exists an integer ℓ with $\ell > t$ such that any state at time t can be determined from knowledge of $\{y_i\}_{i=t}^{\ell}$ and $\{u_i\}_{i=t}^{\ell}$. The periodic subsystem (2.3) is called completely observable if it is observable at all time t.*

(iii) *The backward subsystem (2.4) is observable at time t if there exists an integer s with $s < t$ such that any state at time t can be determined from knowledge of $\{y_i\}_{i=s}^{t}$ and*

$\{u_i\}_{i=s}^{t}$. *The periodic subsystem (2.4) is completely observable if it is observable at all time t.*

Remark. It is easily seen from Definition 3.2 that the periodic discrete-time descriptor system (1.1) is completely observable if and only if both its forward and backward subsystems are completely observable.

We shall state the following Theorems without proofs, which are similar to those of Theorems 3.1 and 3.2.

Theorem 3.3 (Forward Observability). *The following statements are equivalent.*

(a) *The forward subsystem (2.3) is completely observable.*

(b) *For* $t = 0, 1, 2, \cdots, K-1$, *the matrices*

$$\mathcal{O}^f(t) \equiv \begin{bmatrix} C_t^f \\ C_{t+1}^f A_t^f \\ C_{t+2}^f A_{t+1}^f A_t^f \\ \vdots \\ C_{t+n_1K-1}^f \Phi_f(t+n_1K-1,t) \end{bmatrix}$$

have full column rank.

(c) *For* $t = 0, 1, 2, \cdots, K-1$, *and*

$$\mathcal{C}_t^f \equiv \left[(C_t^f)^T,\ (A_t^f)^T(C_{t+1}^f)^T,\ \cdots,\ \Phi_f(t+K-1,t)^T(C_{t+K-1}^f)^T\right]^T,$$

the matrices

$$\begin{bmatrix} \mathcal{C}_t^f \\ \mathcal{C}_t^f \Phi_f(t+K,t) \\ \mathcal{C}_t^f (\Phi_f(t+K,t))^2 \\ \vdots \\ \mathcal{C}_t^f (\Phi_f(t+K,t))^{n_1-1} \end{bmatrix}$$

have full row rank.

(d) *For* $\prod_{i=0}^{K-1} \alpha_i \in \sigma(\Phi_f(K,0))$, *the matrix*

$$V^f(\alpha_0, \cdots, \alpha_{K-1}) \equiv \left[\begin{array}{ccccc} \alpha_0 I & 0 & \cdots & 0 & -A^f_{K-1} \\ -A^f_0 & \alpha_1 I & \ddots & & 0 \\ 0 & -A^f_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -A^f_{K-2} & \alpha_{K-1} I \\ \hline C^f_0 & & & & \\ & C^f_1 & & & \\ & & \ddots & & \\ & & & C^f_{K-2} & \\ & & & & C^f_{K-1} \end{array}\right]$$

has full column rank.

(e) *For* $t = 0, 1, 2, \cdots, K-1$,

$$\Phi_f(t+K, t)x = \lambda x \ \text{ and } \ C^f_i \Phi_f(i,t)x = 0 \ \text{ for } \ i = t, t+1, \cdots, t+K-1$$

imply $x = 0$.

TSING HUA UNIVERSITY

Theorem 3.4 (Backward Observability). *The following statements are equivalent.*

(a) *The backward subsystem (2.4) is completely observable.*

(b) *For* $t = 0, 1, 2, \cdots, K-1$, *the matrices*

$$\mathcal{O}^b(t) \equiv \left[\begin{array}{c} C^b_t \\ C^b_{t-1} E^b_{t-1} \\ C^b_{t-2} E^b_{t-2} E^b_{t-1} \\ \vdots \\ C^b_{t-\nu K+1} \Phi_b(t-\nu K+1, t) \end{array}\right]$$

have full column rank .

(c) *For* $t = 0, 1, 2, \cdots, K-1$, *and*

$$\mathcal{C}_t^b \equiv \left[(C_t^b)^T,\ (E_{t-1}^b)^T(C_{t-1}^b)^T,\ \cdots,\ \Phi_b(t-K+1,t)^T(C_{t-K+1}^b)^T\right]^T,$$

the matrices

$$\begin{bmatrix} \mathcal{C}_t^b \\ \mathcal{C}_t^b N_t \\ \mathcal{C}_t^b N_t^2 \\ \vdots \\ \mathcal{C}_t^b N_t^{\nu-1} \end{bmatrix}$$

have full column rank

(d) *The pair* $(\mathcal{E}_b, \mathcal{C}_b)$ *is observable, where* $\mathcal{E}_b$ *is defined in (3.4) and the matrix* $\mathcal{C}_b \equiv \mathrm{diag}(C_0^b, C_1^b, \cdots, C_{K-1}^b)$.

4 Periodic Reachability and Observability Gramians

It is well known that Gramians play an important role in many applications, such as the model reduction problem [58, 94, 131]. In this Section, the concepts of reachability and observability Gramians are generalized for periodic discrete-time descriptor systems (1.1).

Consider the causal and noncausal reachability matrices given by

$$\mathcal{R}_+(t) \equiv \left[\varphi_{t,t-1}B_{t-1},\ \varphi_{t,t-2}B_{t-2},\cdots,\ \varphi_{t,i}B_i,\cdots\right] \quad (t = 0, 1, \cdots, K-1)$$

and

$$\mathcal{R}_-(t) \equiv \left[\varphi_{t,t}B_t,\ \varphi_{t,t+1}B_{t+1},\cdots,\ \varphi_{t,t+\nu K-1}B_{t+\nu K-1}\right] \quad (t = 0, 1, \cdots, K-1),$$

respectively, with $\varphi_{i,j}$ $(i, j \in \mathbb{Z})$ as defined in (2.6).

Definition 4.1 (Reachability Gramians). *Suppose that the periodic matrix pairs* $\{(E_k, A_k)\}_{k=0}^{K-1}$ *are pd-stable.*

(i) *The causal reachability Gramians of the periodic descriptor system (1.1) are defined by*

$$G_k^{cr} \equiv \mathcal{R}_+(k)\mathcal{R}_+(k)^T = \sum_{i=-\infty}^{k-1} \varphi_{k,i} B_i B_i^T \varphi_{k,i}^T, \quad k = 0, 1, \ldots, K-1.$$

(ii) *The noncausal reachability Gramians of the periodic descriptor system (1.1) are defined by*

$$G_k^{nr} \equiv \mathcal{R}_-(k)\mathcal{R}_-(k)^T = \sum_{i=k}^{k+\nu K-1} \varphi_{k,i} B_i B_i^T \varphi_{k,i}^T, \quad k = 0, 1, \ldots, K-1.$$

(iii) *The reachability Gramians of the periodic descriptor system (1.1) are defined via*

$$G_k^r \equiv G_k^{cr} + G_k^{nr}, \quad k = 0, 1, \ldots, K-1.$$

The causal and noncausal observability matrices are respectively defined by

$$\mathcal{O}_+(t) \equiv \left[\varphi_{t,t-1}^T C_t^T,\ \varphi_{t+1,t-1}^T C_{t+1}^T, \cdots,\ \varphi_{i,t-1}^T C_i^T, \cdots\right]^T \quad (t = 0, 1, \cdots, K-1)$$

and

$$\mathcal{O}_-(t) \equiv \left[\varphi_{t-\nu K,t-1}^T C_{t-\nu K}^T,\ \varphi_{t-\nu K+1,t-1}^T C_{t-\nu K+1}^T, \cdots,\ \varphi_{t-1,t-1}^T C_{t-1}^T\right]^T \quad (t = 0, 1, \cdots, K-1).$$

Definition 4.2 (Observability Gramians). *Suppose that the periodic matrix pairs* $\{(E_k, A_k)\}_{k=0}^{K-1}$ *are pd-stable.*

(i) *The causal observability Gramians of the periodic descriptor system (1.1) are defined by*

$$G_k^{co} \equiv \mathcal{O}_+(k)^T \mathcal{O}_+(k) = \sum_{i=k}^{\infty} \varphi_{i,k-1}^T C_i^T C_i \varphi_{i,k-1}, \quad k = 0, 1, \ldots, K-1.$$

(ii) *The noncausal observability Gramians of the periodic descriptor system (1.1) are defined by*

$$G_k^{no} \equiv \mathcal{O}_-(k)^T \mathcal{O}_-(k) = \sum_{i=k-\nu K}^{k-1} \varphi_{i,k-1}^T C_i^T C_i \varphi_{i,k-1}, \quad k = 0, 1, \ldots, K-1.$$

(iii) *The observability Gramians of the periodic descriptor system (1.1) are defined by*

$$G_k^o \equiv G_k^{co} + G_k^{no}, \quad k = 0, 1, \ldots, K-1.$$

Remarks. (i) The infinite series appeared in the definition of Gramians G_k^{cr} and G_k^{co} converge because of the pd-stability of the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$.

(ii) The Gramians G_k^{cr}, G_k^{nr}, G_k^{co} and G_k^{no} are $n \times n$ symmetric positive semi-definite matrices for all k.

(iii) Definitions 4.1 and 4.2 are natural generalizations of the Gramians defined for descriptor systems with period $K = 1$; see, e.g., [17, 114].

The following theorem indicates that these Gramians of the periodic descriptor system (1.1) satisfy some projected generalized discrete-time periodic Lyapunov equations with special right-hand sides.

Theorem 4.1. *Consider the periodic discrete-time descriptor system (1.1), where the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ are pd-stable.*

(i) *The causal and noncausal reachability Gramians $\{G_k^{cr}\}_{k=0}^{K-1}$ and $\{G_k^{nr}\}_{k=0}^{K-1}$ are the unique symmetric positive semi-definite solutions of the projected GDPLE*

$$\begin{aligned} E_k G_{k+1}^{cr} E_k^T - A_k G_k^{cr} A_k^T &= P_l(k) B_k B_k^T P_l(k)^T, \\ G_k^{cr} &= P_r(k) G_k^{cr} P_r(k)^T, \quad k = 0, 1, 2, \ldots, K-1, \end{aligned} \tag{4.1}$$

and

$$\begin{aligned} E_k G_{k+1}^{nr} E_k^T - A_k G_k^{nr} A_k^T &= -(I - P_l(k)) B_k B_k^T (I - P_l(k))^T, \\ P_r(k) G_k^{nr} &= 0, \quad k = 0, 1, 2, \ldots, K-1, \end{aligned} \tag{4.2}$$

respectively, where $G_K^{cr} \equiv G_0^{cr}$ and $G_K^{nr} \equiv G_0^{nr}$.

(ii) *The causal and noncausal observability Gramians $\{G_k^{co}\}_{k=0}^{K-1}$ and $\{G_k^{no}\}_{k=0}^{K-1}$ are the unique symmetric positive semi-definite solutions of the projected GDPLE*

$$\begin{aligned} E_{k-1}^T G_k^{co} E_{k-1} - A_k^T G_{k+1}^{co} A_k &= P_r(k)^T C_k^T C_k P_r(k), \\ G_k^{co} &= P_l(k-1)^T G_k^{co} P_l(k-1), \quad k = 0, 1, \ldots, K-1, \end{aligned} \tag{4.3}$$

and

$$\begin{aligned} &E_{k-1}^T G_k^{no} E_{k-1} - A_k^T G_{k+1}^{no} A_k = -(I - P_r(k))^T C_k^T C_k (I - P_r(k)), \\ &G_k^{no} P_l(k-1) = 0, \quad k = 0, 1, 2, \ldots, K-1, \end{aligned} \tag{4.4}$$

respectively, where $G_K^{co} \equiv G_0^{co}$, $G_K^{no} \equiv G_0^{no}$, $E_{-1} \equiv E_{K-1}$ *and* $P_l(-1) \equiv P_l(K-1)$.

(iii) *The reachability and observability Gramians* $\{G_k^r\}_{k=0}^{K-1}$ *and* $\{G_k^o\}_{k=0}^{K-1}$ *are the unique symmetric positive semi-definite solutions of the projected GDPLE*

$$\begin{aligned} &E_k G_{k+1}^r E_k^T - A_k G_k^r A_k^T = P_l(k) B_k B_k^T P_l(k)^T - (I - P_l(k)) B_k B_k^T (I - P_l(k))^T, \\ &P_r(k) G_k^r = G_k^r P_r(k)^T, \quad k = 0, 1, 2, \ldots, K-1, \end{aligned} \tag{4.5}$$

and

$$\begin{aligned} &E_{k-1}^T G_k^o E_{k-1} - A_k^T G_{k+1}^o A_k = P_r(k)^T C_k^T C_k P_r(k) - (I - P_r(k))^T C_k^T C_k (I - P_r(k)), \\ &P_l(k-1)^T G_k^o = G_k^o P_l(k-1), \quad k = 0, 1, 2, \ldots, K-1, \end{aligned} \tag{4.6}$$

respectively, where $G_K^r \equiv G_0^r$, $G_K^o \equiv G_0^o$, $E_{-1} \equiv E_{K-1}$ *and* $P_l(-1) \equiv P_l(K-1)$.

Proof. We shall verify only (4.1) here and the other cases can be shown similarly. Rewrite (4.1) into an enlarged Lyapunov equation

$$\mathcal{E}\mathcal{G}\mathcal{E}^T - \mathcal{A}\mathcal{G}\mathcal{A}^T = \mathcal{B}\mathcal{B}^T, \tag{4.7}$$

where

$$\mathcal{E} = \operatorname{diag}(E_0, E_1, \cdots, E_{K-1}), \quad \mathcal{B} = \operatorname{diag}(P_l(0)B_0, P_l(1)B_1, \cdots, P_l(K-1)B_{K-1}),$$

$$\mathcal{A} = \begin{bmatrix} & & & A_0 \\ A_1 & & & \\ & \ddots & & \\ & & A_{K-1} & \end{bmatrix}, \quad \mathcal{G} = \begin{bmatrix} G_1^{cr} & & & \\ & G_2^{cr} & & \\ & & \ddots & \\ & & & G_0^{cr} \end{bmatrix}.$$

Since the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ are pd-stable, the matrix pencil $\lambda\mathcal{E} - \mathcal{A}$ is regular and all its generalized eigenvalues lie inside the unit circle. Then the Lyapunov equation (4.7) has a unique solution and hence the uniqueness of solutions of the projected

GDPLE (4.1) is guaranteed. On the other hand, it can be shown that the causal reachability Gramians G_k^{cr}, $k = 0, 1, \ldots, K-1$, satisfy the projected GDPLE (4.1). Indeed, simple calculation gives that

$$\begin{aligned}
& E_k G_{k+1}^{cr} E_k^T - A_k G_k^{cr} A_k^T \\
= \; & E_k \left(\sum_{i=-\infty}^{k} \varphi_{k+1,i} B_i B_i^T \varphi_{k+1,i}^T \right) E_k^T - A_k \left(\sum_{i=-\infty}^{k-1} \varphi_{k,i} B_i B_i^T \varphi_{k,i}^T \right) A_k^T \\
= \; & E_k Y_{k+1} \left(\sum_{i=-\infty}^{k} \begin{bmatrix} \Phi_f(k+1, i+1) & 0 \\ 0 & 0 \end{bmatrix} X_i B_i B_i^T X_i^T \begin{bmatrix} \Phi_f(k+1, i+1)^T & 0 \\ 0 & 0 \end{bmatrix} \right) Y_{k+1}^T E_k^T \\
& - A_k Y_k \left(\sum_{i=-\infty}^{k-1} \begin{bmatrix} \Phi_f(k, i+1) & 0 \\ 0 & 0 \end{bmatrix} X_i B_i B_i^T X_i^T \begin{bmatrix} \Phi_f(k, i+1)^T & 0 \\ 0 & 0 \end{bmatrix} \right) Y_k^T A_k^T \\
= \; & X_k^{-1} \begin{bmatrix} \sum_{-\infty}^{k} \Phi_f(k+1, i+1) B_i^f (B_i^f)^T \Phi_f(k+1, i+1)^T & 0 \\ 0 & 0 \end{bmatrix} X_k^{-T} \\
& - X_k^{-1} \begin{bmatrix} \sum_{-\infty}^{k-1} \Phi_f(k+1, i+1) B_i^f (B_i^f)^T \Phi_f(k+1, i+1)^T & 0 \\ 0 & 0 \end{bmatrix} X_k^{-T} \\
= \; & X_k^{-1} \begin{bmatrix} B_k^f (B_k^f)^T & 0 \\ 0 & 0 \end{bmatrix} X_k^{-T} = P_l(k) B_k B_k^T P_l(k)^T,
\end{aligned}$$

and

$$\begin{aligned}
& P_r(k) G_k^{cr} P_r(k)^T \\
= \; & Y_k \begin{bmatrix} I_{n_1} & 0 \\ 0 & 0 \end{bmatrix} Y_k^{-1} \left(\sum_{i=-\infty}^{k-1} \varphi_{k,i} B_i B_i^T \varphi_{k,i}^T \right) Y_k^{-T} \begin{bmatrix} I_{n_1} & 0 \\ 0 & 0 \end{bmatrix} Y_k^T \\
= \; & Y_k \begin{bmatrix} \sum_{i=-\infty}^{k-1} \Phi_f(k, i+1) B_i^f (B_i^f)^T \Phi_f(k, i+1)^T & 0 \\ 0 & 0 \end{bmatrix} Y_k^T = G_k^{cr},
\end{aligned}$$

for $k = 0, 1, \ldots, K-1$. Therefore, the causal reachability Gramians $\{G_k^{cr}\}_{k=0}^{K-1}$ are the unique symmetric positive semi-definite solutions of the projected GDPLE (4.1). □

The following theorem shows that complete reachability/observability of the periodic descriptor system (1.1) can be characterized via the reachability/observability Gramians.

Theorem 4.2. *Consider the periodic discrete-time descriptor system (1.1). Assume that the periodic matrix pairs* $\{(E_k, A_k)\}_{k=0}^{K-1}$ *are pd-stable.*

(i) *The periodic descriptor system (1.1) is completely reachable if and only if the reachability Gramians* G_k^r *are positive definite for* $k = 0, 1, 2, \ldots, K-1$.

(ii) *The periodic descriptor system (1.1) is completely observable if and only if the observability Gramians* G_k^o *are positive definite for* $k = 0, 1, 2, \ldots, K-1$.

Proof. Here we shall only prove statement (i) and statement (ii) can be verified similarly. For $k = 0, 1, \ldots, K-1$, pre-multiply (4.5) by X_k and post-multiply (4.5) by X_k^T, it follows that

$$X_k E_k Y_{k+1} \widehat{G}_{k+1}^r Y_{k+1}^T E_k^T X_k^T - X_k A_k Y_k \widehat{G}_k^r Y_k^T A_k^T X_k^T = \begin{bmatrix} B_k^f (B_k^f)^T & 0 \\ 0 & -B_k^b (B_k^b)^T \end{bmatrix}, \quad (4.8)$$

where $\widehat{G}_k^r \equiv Y_k^{-1} G_k^r Y_k^{-T}$.

From Definition 4.1 it is easily seen, for $k = 0, 1, \ldots, K-1$, that

$$\widehat{G}_k^r = Y_k^{-1} G_k^r Y_k^{-T} = \begin{bmatrix} \widehat{G}_{k,1}^{cr} & 0 \\ 0 & \widehat{G}_{k,2}^{nr} \end{bmatrix}, \quad (4.9)$$

with

$$\widehat{G}_{k,1}^{cr} \equiv \sum_{i=-\infty}^{k-1} \Phi_f(k, i+1) B_i^f (B_i^f)^T \Phi_f(k, i+1)^T, \quad \widehat{G}_{k,2}^{nr} \equiv \sum_{i=k}^{k+\nu K-1} \Phi_b(k, i) B_i^b (B_i^b)^T \Phi_b(k, i)^T .$$

Then by (2.1) and (4.9), equations (4.8) are decomposed into two periodic Lyapunov equations, for $k = 0, 1, 2, \ldots, K-1$:

$$\widehat{G}_{k+1,1}^{cr} - A_k^f \widehat{G}_{k,1}^{cr} (A_k^f)^T = B_k^f (B_k^f)^T, \quad (4.10)$$

$$\widehat{G}_{k,2}^{nr} - E_k^b \widehat{G}_{k+1,2}^{nr} (E_k^b)^T = B_k^b (B_k^b)^T. \quad (4.11)$$

Rewrite (4.10) and (4.11) to two enlarged Lyapunov equations:

$$\mathcal{G}_{cr} - \mathcal{A}_f \mathcal{G}_{cr} \mathcal{A}_f^T = \mathcal{B}_f \mathcal{B}_f^T, \quad (4.12)$$

$$\mathcal{G}_{nr} - \mathcal{E}_b \mathcal{G}_{nr} \mathcal{E}_b^T = \mathcal{B}_b \mathcal{B}_b^T, \quad (4.13)$$

where $\mathcal{G}_{cr} = \mathrm{diag}(\widehat{G}^{cr}_{k,1}, \cdots, \widehat{G}^{cr}_{K-1,1}, \widehat{G}^{cr}_{0,1})$, $\mathcal{G}_{nr} = \mathrm{diag}(\widehat{G}^{nr}_{0,2}, \widehat{G}^{nr}_{1,2}, \cdots, \widehat{G}^{nr}_{K-1,2})$, $\mathcal{E}_b$ and $\mathcal{B}_b$ as defined in (3.4), and

$$\mathcal{A}_f = \begin{bmatrix} & & A_0^f \\ A_1^f & & \\ & \ddots & \\ & & A_{K-1}^f \end{bmatrix}, \quad \mathcal{B}_f = \begin{bmatrix} B_0^f & & & \\ & B_1^f & & \\ & & \ddots & \\ & & & B_{K-1}^f \end{bmatrix}.$$

Since the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ are pd-stable and the matrix $\mathcal{E}_b$ is nilpotent with index ν, the pairs $(\mathcal{A}_f, \mathcal{B}_f)$ and $(\mathcal{E}_b, \mathcal{B}_b)$ are reachable if and only if the solutions $\mathcal{G}_{cr}$ and $\mathcal{G}_{nr}$ of Lyapunov equations (4.12), (4.13) are symmetric positive definite. Equivalently, followed from (4.9), all reachability Gramians G_k^r $(k = 0, 1, \ldots, K-1)$ are symmetric positive definite. Moreover, from Theorems 3.1–3.2 and the Remark following Definition 3.1, we know that the periodic descriptor system (1.1) is completely reachable if and only if the pairs $(\mathcal{A}_f, \mathcal{B}_f)$ and $(\mathcal{E}_b, \mathcal{B}_b)$ are reachable. This completes the proof of statement (i). □

5 Numerical Solutions of Projected GDPLEs

In this Section, a numerical method is proposed for the symmetric positive semi-definite solutions of the projected generalized discrete-time periodic Lyapunov equations (4.1) and (4.3), for pd-stable $\{(E_k, A_k)\}_{k=0}^{K-1}$. We first consider the numerical solutions of the GDPLE (4.3).

GDPLE for Observability Gramians G_k^{co}

As $\{(E_k, A_k)\}_{k=0}^{K-1}$ are pd-stable, there exist orthogonal matrices V_k and U_k, with $U_K \equiv U_0$ and for $k = 0, 1, \ldots, K-1$, such that

$$V_k^T E_k U_{k+1} = \begin{bmatrix} E_{k,1} & E_{k,3} \\ 0 & E_{k,2} \end{bmatrix}, \quad V_k^T A_k U_k = \begin{bmatrix} A_{k,1} & A_{k,3} \\ 0 & A_{k,2} \end{bmatrix} \tag{5.1}$$

are upper triangular except $V_0^T A_0 U_0$ is qusi-upper triangular [33, 62]. The matrices $E_{k,1}$ and $A_{k,2}$ are nonsingular, and $E_{k,2}E_{k+1,2}\cdots E_{k+K-1,2}$ are nilpotent for $k = 0, 1, \ldots, K-1$. All finite eigenvalues of the periodic matrix pairs $\{(E_{k,1}, A_{k,1})\}_{k=0}^{K-1}$ lie inside the unit circle and the spectrum of the periodic matrix pairs $\{(E_{k,2}, A_{k,2})\}_{k=0}^{K-1}$ contains only infinite eigenvalues, with

$$\sigma\left(\{(E_{k,1}, A_{k,1})\}_{k=0}^{K-1}\right) \cap \sigma\left(\{(E_{k,2}, A_{k,2})\}_{k=0}^{K-1}\right) = \emptyset. \tag{5.2}$$

Computationally, these matrix decompositions can be accomplished via the periodic QZ algorithm (PQZ) with reordering strategies.

Notice that

$$\begin{bmatrix} I & Z_k \\ 0 & I \end{bmatrix} \begin{bmatrix} E_{k,1} & E_{k,3} \\ 0 & E_{k,2} \end{bmatrix} \begin{bmatrix} I & -W_{k+1} \\ 0 & I \end{bmatrix} = \begin{bmatrix} E_{k,1} & 0 \\ 0 & E_{k,2} \end{bmatrix}, \tag{5.3}$$

$$\begin{bmatrix} I & Z_k \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{k,1} & A_{k,3} \\ 0 & A_{k,2} \end{bmatrix} \begin{bmatrix} I & -W_k \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{k,1} & 0 \\ 0 & A_{k,2} \end{bmatrix}, \tag{5.4}$$

if the matrices Z_k and W_k, with $W_K = W_0$ and for $k = 0, 1, \ldots, K-1$, satisfy the generalized periodic Sylvester equations

$$\begin{aligned} E_{k,1}W_{k+1} - Z_k E_{k,2} &= E_{k,3}, \\ A_{k,1}W_k - Z_k A_{k,2} &= A_{k,3}. \end{aligned} \tag{5.5}$$

From condition (5.2), the generalized periodic Sylvester equations (5.5) have unique solutions Z_k and W_k. Therefore, the nonsingular matrices X_k, Y_k in (2.1) satisfy

$$X_k = \begin{bmatrix} I & Z_k \\ 0 & I \end{bmatrix} V_k^T, \quad Y_k = U_k \begin{bmatrix} I & -W_k \\ 0 & I \end{bmatrix},$$

and the right and left spectral projections $P_r(k)$, $P_l(k)$ are given as

$$P_l(k) = V_k \begin{bmatrix} I & Z_k \\ 0 & 0 \end{bmatrix} V_k^T, \quad P_r(k) = U_k \begin{bmatrix} I & W_k \\ 0 & 0 \end{bmatrix} U_k^T. \tag{5.6}$$

Let, for $k = 0, 1, \ldots, K-1$,

$$V_{k-1}^T G_k^{co} V_{k-1} = \begin{bmatrix} G_{k,1}^{co} & G_{k,3}^{co} \\ (G_{k,3}^{co})^T & G_{k,2}^{co} \end{bmatrix}, \quad C_k U_k = \begin{bmatrix} C_{k,1}, & C_{k,2} \end{bmatrix}. \tag{5.7}$$

Substituting (5.1), (5.6) and (5.7) into the projected GDPLE (4.3), for $k = 0, 1, \ldots, K-1$, we have

$$E_{k-1,1}^T G_{k,1}^{co} E_{k-1,1} - A_{k,1}^T G_{k+1,1}^{co} A_{k,1} = C_{k,1}^T C_{k,1}, \tag{5.8}$$

$$E_{k-1,1}^T G_{k,1}^{co} E_{k-1,3} + E_{k-1,1}^T G_{k,3}^{co} E_{k-1,2} - A_{k,1}^T G_{k+1,1}^{co} A_{k,3} - A_{k,1}^T G_{k+1,3}^{co} A_{k,2} = C_{k,1}^T C_{k,1} W_k, \tag{5.9}$$

$$E_{k-1,3}^T G_{k,1}^{co} E_{k-1,3} + E_{k-1,3}^T G_{k,3}^{co} E_{k-1,2} + E_{k-1,2}^T (G_{k,3}^{co})^T E_{k-1,3} + E_{k-1,2}^T G_{k,2}^{co} E_{k-1,2} -$$
$$A_{k,3}^T G_{k+1,1}^{co} A_{k,3} - A_{k,3}^T G_{k+1,3}^{co} A_{k,2} - A_{k,2}^T (G_{k+1,3}^{co})^T A_{k,3} - A_{k,2}^T G_{k+1,2}^{co} A_{k,2} = W_k^T C_{k,1}^T C_{k,1} W_k. \tag{5.10}$$

Again from the pd-stability of $\{(E_{k,1}, A_{k,1})\}_{k=0}^{K-1}$, the generalized discrete-time periodic Lyapunov equations (5.8) have unique symmetric positive semi-definite solutions $G_{k,1}^{co}$. Furthermore, it follows from (5.5) that (5.9) can be rearranged as

$$E_{k-1,1}^T (G_{k,3}^{co} - G_{k,1}^{co} Z_{k-1}) E_{k-1,2} - A_{k,1}^T (G_{k+1,3}^{co} - G_{k+1,1}^{co} Z_k) A_{k,2} = 0. \tag{5.11}$$

Again, from (5.2), we deduce that

$$G_{k,3}^{co} = G_{k,1}^{co} Z_{k-1}, \quad k = 0, 1, \ldots, K-1. \tag{5.12}$$

From (5.5), (5.8) and (5.12), (5.10) can be rewritten as

$$E_{k-1,2}^T (G_{k,2}^{co} - Z_{k-1}^T G_{k,1}^{co} Z_{k-1}) E_{k-1,2} - A_{k,2}^T (G_{k+1,2}^{co} - Z_k^T G_{k+1,1}^{co} Z_k) A_{k,2} = 0. \tag{5.13}$$

Now, since the periodic matrix pairs $\{(E_{k,2}, A_{k,2})\}_{k=0}^{K-1}$ have only infinite eigenvalues, we then have

$$G_{k,2}^{co} = Z_{k-1}^T G_{k,1}^{co} Z_{k-1}, \quad k = 0, 1, \ldots, K-1. \tag{5.14}$$

Therefore, the solutions of the projected GDPLE (4.3) have the form

$$G_k^{co} = V_{k-1} \begin{bmatrix} G_{k,1}^{co} & G_{k,1}^{co} Z_{k-1} \\ Z_{k-1}^T G_{k,1}^{co} & Z_{k-1}^T G_{k,1}^{co} Z_{k-1} \end{bmatrix} V_{k-1}^T, \quad k = 0, 1, \ldots, K-1, \tag{5.15}$$

where the matrices $G_{k,1}^{co}$ are the unique symmetric positive semi-definite solutions of the generalized periodic Lyapunov equations (5.8). Moreover, from (5.6) and (5.15) they also satisfy $P_l(k-1)^T G_k^{co} P_l(k-1) = G_k^{co}$.

In many applications it is necessary to have the Cholesky factors of the solutions of the Lyapunov equations rather the solutions itself [78]. In particular, these full-ranked factors are useful for computing numerically the Hankel singular values (see next Section). If $L_{k,1}$ denotes a Cholesky factor of each matrix $G^{co}_{k,1}$, i.e., $G^{co}_{k,1} = L^T_{k,1}L_{k,1}$, then we compute the QR factorization

$$L_{k,1} = Q_{k,L}\begin{bmatrix} T_{k,L} \\ 0 \end{bmatrix},$$

where $Q_{k,L}$ is orthogonal and $T_{k,L}$ has full row rank, for $k = 0, 1, \ldots, K-1$. The full-ranked factorizations of the solutions G^{co}_k, for $k = 0, 1, \ldots, K-1$, are given by

$$\begin{aligned} G^{co}_k &= V_{k-1}\begin{bmatrix} L^T_{k,1} \\ Z^T_{k-1}L^T_{k,1} \end{bmatrix}\begin{bmatrix} L_{k,1}, & L_{k,1}Z_{k-1} \end{bmatrix}V^T_{k-1} \\ &= V_{k-1}\begin{bmatrix} T^T_{k,L} \\ Z^T_{k-1}T^T_{k,L} \end{bmatrix}\begin{bmatrix} T_{k,L}, & T_{k,L}Z_{k-1} \end{bmatrix}V^T_{k-1} \\ &\equiv L^T_k L_k, \end{aligned}$$

where $L_k \equiv \begin{bmatrix} T_{k,L}, & T_{k,L}Z_{k-1} \end{bmatrix}V^T_{k-1}$ has full row-rank.

GDPLE for Reachability Gramians G^{cr}_k

Similarly for the projected GDPLE (4.1), for $k = 0, 1, \ldots, K-1$, we let

$$U^T_k G^{cr}_k U_k = \begin{bmatrix} G^{cr}_{k,1} & G^{cr}_{k,3} \\ (G^{cr}_{k,3})^T & G^{cr}_{k,2} \end{bmatrix}, \quad V^T_k B_k = \begin{bmatrix} B_{k,1} \\ B_{k,2} \end{bmatrix}. \tag{5.16}$$

Substituting (5.1), (5.6) and (5.16) into the projected GDPLE (4.1), we then have

$$\begin{aligned} E_{k,1}G^{cr}_{k+1,1}E^T_{k,1} - A_{k,1}G^{cr}_{k,1}A^T_{k,1} &= -E_{k,1}G^{cr}_{k+1,3}E^T_{k,3} - E_{k,3}(G^{cr}_{k+1,3})^T E^T_{k,1} - E_{k,3}G^{cr}_{k+1,2}E^T_{k,3} \\ &\quad + A_{k,1}G^{cr}_{k,3}A^T_{k,3} + A_{k,3}(G^{cr}_{k,3})^T A^T_{k,1} + A_{k,3}G^{cr}_{k,2}A^T_{k,3} \\ &\quad + (B_{k,1} + Z_k B_{k,2})(B_{k,1} + Z_k B_{k,2})^T, \end{aligned} \tag{5.17}$$

$$E_{k,1}G^{cr}_{k+1,3}E^T_{k,2} - A_{k,1}G^{cr}_{k,3}A^T_{k,2} = -E_{k,3}G^{cr}_{k+1,2}E^T_{k,2} + A_{k,3}G^{cr}_{k,2}A^T_{k,2}, \tag{5.18}$$

$$E_{k,2}G^{cr}_{k+1,2}E^T_{k,2} - A_{k,2}G^{cr}_{k,2}A^T_{k,2} = 0, \quad k = 0, 1, \ldots, K-1. \tag{5.19}$$

Since the periodic matrix pairs $\{(E_{k,2}, A_{k,2})\}_{k=0}^{K-1}$ have only infinite eigenvalues, it follows from (5.19) that

$$G_{k,2}^{cr} = 0, \quad k = 0, 1, \ldots, K-1. \tag{5.20}$$

Furthermore, (5.18) can be simplified to

$$E_{k,1} G_{k+1,3}^{cr} E_{k,2}^T - A_{k,1} G_{k,3}^{cr} A_{k,2}^T = 0. \tag{5.21}$$

Then from (5.2), we have

$$G_{k,3}^{cr} = 0, \quad k = 0, 1, \ldots, K-1. \tag{5.22}$$

From (5.20) and (5.22), (5.17) can be rewritten as

$$E_{k,1} G_{k+1,1}^{cr} E_{k,1}^T - A_{k,1} G_{k,1}^{cr} A_{k,1}^T = (B_{k,1} + Z_k B_{k,2})(B_{k,1} + Z_k B_{k,2})^T. \tag{5.23}$$

Therefore, the solutions of the projected GDPLE (4.1) have the form

$$G_k^{cr} = U_k \begin{bmatrix} G_{k,1}^{cr} & 0 \\ 0 & 0 \end{bmatrix} U_k^T, \quad k = 0, 1, \ldots, K-1, \tag{5.24}$$

where the matrices $G_{k,1}^{cr}$ are the unique symmetric positive semi-definite solutions of the generalized periodic Lyapunov equations (5.23). Moreover, from (5.6) and (5.24) they also satisfy $P_r(k) G_k^{cr} P_r(k)^T = G_k^{cr}$.

If $R_{k,1}$ denotes a Cholesky factor of each matrix $G_{k,1}^{cr}$, i.e., $G_{k,1}^{cr} = R_{k,1} R_{k,1}^T$, then we compute the QR factorization

$$R_{k,1}^T = Q_{k,R} \begin{bmatrix} T_{k,R}^T \\ 0 \end{bmatrix},$$

where $Q_{k,R}$ is orthogonal and $T_{k,R}$ has full column-rank. The full-ranked factorizations of the solutions G_k^{cr} are given by

$$\begin{aligned} G_k^{cr} &= U_k \begin{bmatrix} R_{k,1} \\ 0 \end{bmatrix} \begin{bmatrix} R_{k,1}^T, & 0 \end{bmatrix} U_k^T \\ &= U_k \begin{bmatrix} T_{k,R} \\ 0 \end{bmatrix} \begin{bmatrix} T_{k,R}^T, & 0 \end{bmatrix} U_k^T \\ &\equiv R_k R_k^T, \end{aligned}$$

where $R_k^T \equiv \left[\begin{array}{cc} T_{k,R}^T, & 0 \end{array}\right] U_k^T$ has full row-rank for $k = 0, 1, \ldots, K-1$.

Algorithm GDPLEs

We now summarize the main steps for computing the full-ranked Cholesky factors of the causal Gramians, via the solution of the GDPLEs (4.1) and (4.3). For simplicity in Algorithm 5.1, we shall ignore the obvious qualification for k, i.e., $k = 0, 1, \cdots, K-1$.

Algorithm 5.1 (GDPLEs)

Input: System matrices (E_k, A_k, B_k, C_k), with $\{(E_k, A_k)\}_{k=0}^{K-1}$ being pd-stable.

Output: Full-ranked Cholesky factors R_k and L_k $(k = 0, 1, \ldots, K-1)$, where
$G_k^{cr} = R_k R_k^T$ and $G_k^{co} = L_k^T L_k$.

Step 1. Use the PQZ algorithm [33, 62] to compute orthogonal matrices V_k and U_k, with $U_K \equiv U_0$, such that

$$V_k^T E_k U_{k+1} = \begin{bmatrix} E_{k,1} & E_{k,3} \\ 0 & E_{k,2} \end{bmatrix}, \quad V_k^T A_k U_k = \begin{bmatrix} A_{k,1} & A_{k,3} \\ 0 & A_{k,2} \end{bmatrix}$$

are upper triangular except $V_0^T A_0 U_0$ is quasi-upper triangular. The matrices $E_{k,1}$ and $A_{k,2}$ are nonsingular, and $E_{k,2}E_{k+1,2}\cdots E_{k+K-1,2}$ are nilpotent.

Step 2. Use the Cyclic Schur and Hessenberg-Schur methods [43] to compute the solutions of the generalized periodic Sylvester equations

$$E_{k,1} W_{k+1} - Z_k E_{k,2} = E_{k,3},$$
$$A_{k,1} W_k - Z_k A_{k,2} = A_{k,3}.$$

Step 3. Compute the matrices

$$V_k^T B_k = \begin{bmatrix} B_{k,1} \\ B_{k,2} \end{bmatrix}, \quad C_k U_k = \left[\begin{array}{cc} C_{k,1}, & C_{k,2} \end{array}\right].$$

Step 4. Compute the Cholesky factors $R_{k,1}$ and $L_{k,1}$ of the solutions $G_{k,1}^{cr} = R_{k,1}R_{k,1}^T$ and $G_{k,1}^{co} = L_{k,1}^T L_{k,1}$ of the generalized discrete-time periodic Lyapunov equations

$$E_{k,1}G_{k+1,1}^{cr}E_{k,1}^T - A_{k,1}G_{k,1}^{cr}A_{k,1}^T = (B_{k,1} + Z_k B_{k,2})(B_{k,1} + Z_k B_{k,2})^T,$$
$$E_{k-1,1}^T G_{k,1}^{co} E_{k-1,1} - A_{k,1}^T G_{k+1,1}^{co} A_{k,1} = C_{k,1}^T C_{k,1}.$$

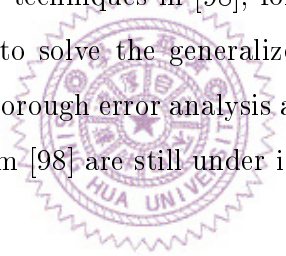
Step 5. Compute the QR factorizations

$$R_{k,1}^T = Q_{k,R}\begin{bmatrix} T_{k,R}^T \\ 0 \end{bmatrix}, \quad L_{k,1} = Q_{k,L}\begin{bmatrix} T_{k,L} \\ 0 \end{bmatrix}.$$

Step 6. Compute the full-ranked Cholesky factors

$$R_k = U_k \begin{bmatrix} T_{k,R} \\ 0 \end{bmatrix}, \quad L_k = \begin{bmatrix} T_{k,L}, & T_{k,L}Z_{k-1} \end{bmatrix} V_{k-1}^T.$$

Remark. One can extend the techniques in [98], for the numerical solution of the generalized Lyapunov equations, to solve the generalized discrete-time periodic Lyapunov equations given in Step 4. A thorough error analysis and practical implementation details for the algorithm extended from [98] are still under investigation.



A Numerical Example

We shall illustrate the feasibility and reliability of the proposed algorithm with an example. All computations were performed in MATLAB/version 6.5 on a PC with an Intel Pentium-III processor at 866 MHz, with 768 MB RAM, using IEEE double-precision floating-point arithmetic. The machine precision is approximately 2.22×10^{-16}.

For approximate solutions $\widetilde{X}_k$ of the projected generalized discrete-time periodic Lyapunov equations (4.1) and (4.3), we compute the relative residuals defined by

$$\gamma_k^{cr} = \frac{\|E_k\widetilde{X}_{k+1}E_k^T - A_k\widetilde{X}_k A_k^T - P_l(k)B_k B_k^T P_l(k)^T\|_2}{\|\widetilde{X}_k\|_2},$$
$$\gamma_k^{co} = \frac{\|E_{k-1}^T\widetilde{X}_k E_{k-1} - A_k^T\widetilde{X}_{k+1}A_k - P_r(k)^T C_k^T C_k P_r(k)\|_2}{\|\widetilde{X}_k\|_2}.$$

Example 1. We consider a periodic discrete-time descriptor system (1.1) with $n = 10$, $m = 2$, $p = 3$ and period $K = 3$. For $k = 0, 1, 2$, we have

$$E_k^{(0)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & c_1 & s_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -s_1 & c_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_1 & s_1 & 1 & 0 & c_2 & s_2 & 0 & 0 & 0 \\ 0 & -s_1 & c_1 & 0 & 1 & -s_2 & c_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_2 & s_2 & 1 & 0 & c_3 & s_3 & 0 \\ 0 & 0 & 0 & -s_2 & c_2 & 0 & 1 & -s_3 & c_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & c_3 & s_3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -s_3 & c_3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_k^{(0)} = \operatorname{diag}(1.01, A_{01}, A_{02}, A_{03}, A_{04}, 1.001), \quad \theta_k := 2\pi k/K,$$

$$B_k^T = \begin{bmatrix} 4 & -1 & 3 & 5 & 0 & -2 & 0 & 8 & 1 & 0 \\ 1 & 1 & s_1 + 1 & -2 & 1 & 0 & 0 & -3 & 0 & 1 \end{bmatrix},$$

$$C_k = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.05 + c_1 & 0 & 0 \end{bmatrix},$$

where

$$c_1 = \cos(\theta_k), \quad c_2 = 0.2c_1, \quad c_3 = 0.6c_1,$$

$$s_1 = \sin(\theta_k), \quad s_2 = 0.2s_1, \quad s_3 = 0.6s_1,$$

$$A_{01} = \begin{bmatrix} r_1\cos(\pi/3) & r_1\sin(\pi/3) \\ -r_1\sin(\pi/3) & r_1\cos(\pi/3) \end{bmatrix}, \quad A_{02} = \begin{bmatrix} r_2\cos(7\pi/5) & r_2\sin(7\pi/5) \\ -r_2\sin(7\pi/5) & r_2\cos(7\pi/5) \end{bmatrix},$$

$$A_{03} = \begin{bmatrix} r_3\cos(\pi/4) & r_3\sin(\pi/4) \\ -r_3\sin(\pi/4) & r_3\cos(\pi/4) \end{bmatrix}, \quad A_{04} = \begin{bmatrix} r_4\cos(\pi/10) & r_4\sin(\pi/10) \\ -r_4\sin(\pi/10) & r_4\cos(\pi/10) \end{bmatrix},$$

and

$$r_1 = 0.5, \quad r_2 = 0.05, \quad r_3 = -0.02, \quad r_4 = 0.12.$$

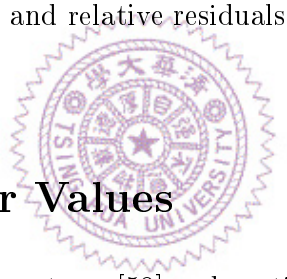
We define a Householder transformation $V = I - 2uu^T$ with $u = [\,1, 1, \cdots, 1, 1\,]^T/\sqrt{10} \in \mathbb{R}^{10}$, and the K-periodic system matrices (E_k, A_k, B_k, C_k) are given by

$$E_k \equiv V^T E_k^{(0)} V, \quad A_k \equiv V^T A_k^{(0)} V, \quad k = 0, 1, 2.$$

The computed open-loop spectrum of the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ consists of two infinite eigenvalues and four pairs of complex conjugate finite eigenvalues lying inside the unit circle. Thus, the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ are pd-stable with $n_1 = 8$ and $n_2 = 2$. Accurate numerical results were produced by the proposed algorithm, as shown in Table 1.

k	$\|G_k^{cr}\|_2$	γ_k^{cr}	$\|G_k^{co}\|_2$	γ_k^{co}
0	8.30×10^4	2.17×10^{-16}	1.14×10^3	1.39×10^{-16}
1	7.11×10^3	3.11×10^{-16}	9.70×10^0	4.17×10^{-15}
2	5.82×10^2	6.73×10^{-16}	9.74×10^1	9.18×10^{-15}

Table 1: Norms and relative residuals of causal Gramians.



6 Hankel Singular Values

Similar to standard state space systems [58] and continuous-time descriptor systems [110, 113], the controllability and observability Gramians can be used to define Hankel singular values for the periodic descriptor systems (1.1), which are of great importance in the model reduction problem via the balanced truncation method.

For the discrete-time descriptor systems, the causal and noncausal Hankel singular values are defined via the nonnegative eigenvalues of the matrices $\mathcal{G}_{dcc}E^T\mathcal{G}_{dco}E$ and $\mathcal{G}_{dnc}A^T\mathcal{G}_{dno}A$. Here $\mathcal{G}_{dcc}$, $\mathcal{G}_{dnc}$, $\mathcal{G}_{dco}$ and $\mathcal{G}_{dno}$ denote the causal/noncausal reachability Gramians and the causal/noncausal observability Gramians, respectively [114].

Lemma 6.1. *Let the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ be pd-stable. Then the matrices $\mathbf{H}_k^c \equiv G_k^{cr}E_{k-1}^T G_k^{co} E_{k-1}$ and $\mathbf{H}_k^{nc} \equiv G_k^{nr}A_k^T G_{k+1}^{no} A_k$, $k = 0, 1, 2, \ldots, K-1$, have real and nonnegative eigenvalues.*

Proof. From Definitions 4.1, 4.2 and (2.6) and for $k = 0, 1, 2, \ldots, K-1$, we have

$$\mathbf{H}_k^c = Y_k \begin{bmatrix} \widehat{G}_{k,1}^{cr} \widehat{G}_{k,1}^{co} & 0 \\ 0 & 0 \end{bmatrix} Y_k^{-1},$$

where

$$\widehat{G}_{k,1}^{cr} \equiv \sum_{i=-\infty}^{k-1} \Phi_f(k, i+1) B_i^f (B_i^f)^T \Phi_f(k, i+1)^T, \quad \widehat{G}_{k,1}^{co} \equiv \sum_{i=k}^{\infty} \Phi_f(i, k) (C_i^f)^T C_i^f \Phi_f(i, k).$$

Since the $n_1 \times n_1$ matrices $\widehat{G}_{k,1}^{cr}$ and $\widehat{G}_{k,1}^{co}$ are symmetric positive semi-definite, it follows that $\mathbf{H}_k^c$ have real and nonnegative eigenvalues. Similarly, it can be shown that $\mathbf{H}_k^{nc}$ also share the same property. □

Notice that, in the proof of Lemma 6.1, the matrices $\mathbf{H}_k^c$ and $\mathbf{H}_k^{nc}$ have at least n_2 and n_1 zero eigenvalues, respectively. Hence, we have the following definition of Hankel singular values for the periodic descriptor system (1.1).

Definition 6.1. *Suppose that the periodic matrix pairs $\{(E_k, A_k)\}_{k=0}^{K-1}$ are pd-stable and let n_1, n_2 be the dimensions of the periodic deflating subspaces of $\{(E_k, A_k)\}_{k=0}^{K-1}$ corresponding respectively to the finite and infinite eigenvalues.*

(i) *For $k = 0, 1, \ldots, K-1$, the square roots of the largest n_1 eigenvalues of the matrices $\mathbf{H}_k^c$, denoted by $\zeta_{k,j}$, are called the causal Hankel singular values of the periodic descriptor system (1.1).*

(ii) *For $k = 0, 1, \ldots, K-1$, the square roots of the largest n_2 eigenvalues of the matrices $\mathbf{H}_k^{nc}$, denoted by $\theta_{k,j}$, are called the noncausal Hankel singular values of the periodic descriptor system (1.1).*

Remarks. (i) When $K = 1$, the causal and noncausal Hankel singular values defined in Definition 6.1 coincide with those for discrete-time descriptor systems (see [114] and references therein). For $E_k = I$, the causal Hankel singular values are the classical Hankel singular values of linear periodic discrete-time systems [124].

(ii) As in the case of descriptor systems, the causal and noncausal Hankel singular values of the periodic descriptor system (1.1) are invariant under system equivalence transformations.

From Theorem 4.2 and Lemma 6.1 we obtain the following result.

Corollary 6.2. *Consider the periodic discrete-time descriptor system (1.1), where the periodic matrix pairs* $\{(E_k, A_k)\}_{k=0}^{K-1}$ *are pd-stable. The following statements are equivalent.*

(a) *The periodic descriptor system (1.1) is completely reachable and completely observable.*

(b) *For* $k = 0, 1, 2, \ldots, K-1$, *we have*

$$\begin{aligned} \operatorname{rank}(G_k^{cr}) &= \operatorname{rank}(G_k^{co}) = \operatorname{rank}(\mathbf{H}_k^c) = n_1, \\ \operatorname{rank}(G_k^{nr}) &= \operatorname{rank}(G_k^{no}) = \operatorname{rank}(\mathbf{H}_k^{nc}) = n_2. \end{aligned}$$

(c) *The causal and noncausal Hankel singular values of (1.1) are nonzero.*

For pd-stable $\{(E_k, A_k)\}_{k=0}^{K-1}$, the causal and noncausal reachability and observability Gramians are symmetric and positive semi-definite. Thus, there exist full-ranked factorizations

$$\begin{aligned} G_k^{cr} &= R_k R_k^T, \quad G_k^{co} = L_k^T L_k, \\ G_k^{nr} &= \widetilde{R}_k \widetilde{R}_k^T, \quad G_k^{co} = \widetilde{L}_k^T \widetilde{L}_k, \end{aligned} \tag{6.1}$$

where the matrices R_k, L_k^T, $\tilde{R}_k$ and $\tilde{L}_k^T$ are of full column-rank. The connections between the causal/noncausal Hankel singular values and the singular values of the matrices $L_k E_{k-1} R_k$ and $\widetilde{L}_{k+1} A_k \widetilde{R}_k$ are considered in the following Lemma.

Lemma 6.3. *For the periodic descriptor system (1.1), where the periodic matrix pairs* $\{(E_k, A_k)\}_{k=0}^{K-1}$ *are pd-stable. Suppose that the causal and noncausal Gramians of (1.1) have the full-ranked factorizations defined as in (6.1). Then for* $k = 0, 1, 2, \ldots, K-1$, *the nonzero causal Hankel singular values are the nonzero singular values of the matrices* $L_k E_{k-1} R_k$, *while the nonzero noncausal Hankel singular values are the nonzero singular values of the matrices* $\widetilde{L}_{k+1} A_k \widetilde{R}_k$.

Proof. Notice that for $k = 0, 1, \ldots, K-1$, we have

$$\begin{aligned} \zeta_{k,j}^2 &= \lambda_j(R_k R_k^T E_{k-1}^T L_k^T L_k E_{k-1}) = \lambda_j(R_k^T E_{k-1}^T L_k^T L_k E_{k-1} R_k) = \sigma_j^2(L_k E_{k-1} R_k), \\ \theta_{k,j}^2 &= \lambda_j(\widetilde{R}_k \widetilde{R}_k^T A_k^T \widetilde{L}_{k+1}^T \widetilde{L}_{k+1} A_k) = \lambda_j(\widetilde{R}_k^T A_k^T \widetilde{L}_{k+1}^T \widetilde{L}_{k+1} A_k \widetilde{R}_k) = \sigma_j^2(\widetilde{L}_{k+1} A_k \widetilde{R}_k), \end{aligned}$$

where $\lambda_j(\cdot)$ and $\sigma_j(\cdot)$ denote, respectively, the eigenvalues and singular values of the corresponding matrices. □

7 Balanced Realization

It is well known [58] that for any minimal realization (A, B, C) of a stable continuous-time or discrete-time system, there exists a transformation such that the controllability and observability Gramians for the transformed realization equal to some diagonal matrix. Such a realization is called a(n) (internally) balanced realization. Recently, the issues of balanced realization and model reduction via the balanced truncation method are discussed for continuous-time descriptor systems [110, 113] and asymptotically stable linear discrete-time periodic systems [123, 124]. In this Section the problem of balanced realization is generalized for periodic descriptor systems. We shall assume that the periodic descriptor system (1.1) is completely reachable/observable with $\{(E_k, A_k)\}_{k=0}^{K-1}$ being pd-stable.

Definition 7.1. *A realization (E_k, A_k, B_k, C_k) of the periodic descriptor system (1.1) is called balanced if*

$$G_k^{cr} = G_k^{co} = \begin{bmatrix} D_{k,1} & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad G_k^{nr} = G_{k+1}^{no} = \begin{bmatrix} 0 & 0 \\ 0 & D_{k,2} \end{bmatrix},$$

where $D_{k,1}$ and $D_{k,2}$ are diagonal matrices for $k = 0, 1, \ldots, K-1$.

We shall show that for a realization (E_k, A_k, B_k, C_k) of the periodic descriptor system (1.1), there exist nonsingular periodic matrices S_k and T_k $(k = 0, 1, \ldots, K-1)$ with $T_K \equiv T_0$, such that the transformed realization

$$(\widehat{E}_k, \widehat{A}_k, \widehat{B}_k, \widehat{C}_k) \equiv (S_k^T E_k T_{k+1}, S_k^T A_k T_k, S_k^T B_k, C_k T_k) \tag{7.1}$$

is balanced.

Consider the full-ranked factorizations (6.1) of the causal and noncausal reachability and observability Gramians. For $k = 0, 1, \cdots, K-1$, let

$$L_k E_{k-1} R_k = U_k \Sigma_k V_k^T, \quad \widetilde{L}_{k+1} A_k \widetilde{R}_k = \widetilde{U}_k \Theta_k \widetilde{V}_k^T, \tag{7.2}$$

be the singular value decompositions [59] of $L_k E_{k-1} R_k$ and $\widetilde{L}_{k+1} A_k \widetilde{R}_k$. Here U_k, V_k, $\widetilde{U}_k$, $\widetilde{V}_k$ are orthogonal, and Σ_k and Θ_k are diagonal and nonsingular. From Corollary 6.2 and Lemma 6.3, we have $\Sigma_k = \mathrm{diag}(\zeta_{k,1}, \cdots, \zeta_{k,n_1}) > 0$ and $\Theta_k = \mathrm{diag}(\theta_{k,1}, \cdots, \theta_{k,n_2}) > 0$. Furthermore, it is easily seen from Theorem 4.1 and (2.5) that

$$\begin{aligned}
&G_k^{cr} = P_r(k) G_k^{cr} P_r(k)^T, \quad G_k^{co} = P_l(k-1)^T G_k^{co} P_l(k-1), \\
&P_r(k) G_k^{nr} = 0, \quad G_k^{no} P_l(k-1) = 0, \\
&E_{k-1} P_r(k) = P_l(k-1) E_{k-1}, \quad A_k P_r(k) = P_l(k) A_k.
\end{aligned}$$

Simple calculations then yield $G_k^{no} E_{k-1} G_k^{cr} = G_k^{co} E_{k-1} G_k^{nr} = G_{k+1}^{no} A_k G_k^{cr} = G_{k+1}^{co} A_k G_k^{nr} = 0$. Hence, for $k = 0, 1, \ldots, K-1$, we have

$$\widetilde{L}_k E_{k-1} R_k = L_k E_{k-1} \widetilde{R}_k = \widetilde{L}_{k+1} A_k R_k = L_{k+1} A_k \widetilde{R}_k = 0. \tag{7.3}$$

Now for $k = 0, 1, \ldots, K-1$, consider the $n \times n$ matrices

$$S_k = \begin{bmatrix} L_{k+1}^T U_{k+1} \Sigma_{k+1}^{-1/2}, & \widetilde{L}_{k+1}^T \widetilde{U}_k \Theta_k^{-1/2} \end{bmatrix}, \quad \check{S}_k = \begin{bmatrix} E_k R_{k+1} V_{k+1} \Sigma_{k+1}^{-1/2}, & A_k \widetilde{R}_k \widetilde{V}_k \Theta_k^{-1/2} \end{bmatrix},$$

It follows from (7.2) and (7.3) that

$$S_k^T \check{S}_k = \begin{bmatrix} \Sigma_{k+1}^{-1/2} U_{k+1}^T L_{k+1} E_k R_{k+1} V_{k+1} \Sigma_{k+1}^{-1/2} & \Sigma_{k+1}^{-1/2} U_{k+1}^T L_{k+1} A_k \widetilde{R}_k \widetilde{V}_k \Theta_k^{-1/2} \\ \Theta_k^{-1/2} \widetilde{U}_k^T \widetilde{L}_{k+1} E_k R_{k+1} V_{k+1} \Sigma_{k+1}^{-1/2} & \Theta_k^{-1/2} \widetilde{U}_k^T \widetilde{L}_{k+1} A_k \widetilde{R}_k \widetilde{V}_k \Theta_k^{-1/2} \end{bmatrix} = I_n,$$

i.e., the matrices S_k and $\check{S}_k$ are nonsingular and $S_k^{-1} = \check{S}_k^T$. Similarly, it can be shown that the matrices

$$T_k = \begin{bmatrix} R_k V_k \Sigma_k^{-1/2}, & \widetilde{R}_k \widetilde{V}_k \Theta_k^{-1/2} \end{bmatrix}, \quad \check{T}_k = \begin{bmatrix} E_{k-1}^T L_k^T U_k \Sigma_k^{-1/2}, & A_k^T \widetilde{L}_{k+1}^T \widetilde{U}_k \Theta_k^{-1/2} \end{bmatrix}$$

are also nonsingular and $T_k^{-1} = \check{T}_k^T$. Therefore, with the transformation matrices S_k and T_k defined above and (7.3), the causal reachability and observability Gramians of the

transformed periodic descriptor system (7.1) become

$$\begin{aligned}\widehat{G}_k^{cr} &\equiv T_k^{-1} G_k^{cr} T_k^{-T} = \check{T}_k^T G_k^{cr} \check{T}_k \\ &= \begin{bmatrix} \Sigma_k^{-1/2} U_k^T L_k E_{k-1} R_k R_k^T E_{k-1}^T L_k^T U_k \Sigma_k^{-1/2} & \Sigma_k^{-1/2} U_k^T L_k E_{k-1} R_k R_k^T A_k^T \widetilde{L}_{k+1}^T \widetilde{U}_k \Theta_k^{-1/2} \\ \Theta_k^{-1/2} \widetilde{U}_k^T \widetilde{L}_{k+1} A_k R_k R_k^T E_{k-1}^T L_k^T U_k \Sigma_k^{-1/2} & \Theta_k^{-1/2} \widetilde{U}_k^T \widetilde{L}_{k+1} A_k R_k R_k^T A_k^T \widetilde{L}_{k+1}^T \widetilde{U}_k \Theta_k^{-1/2} \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_k & 0 \\ 0 & 0 \end{bmatrix},\end{aligned}$$

and

$$\begin{aligned}&\widehat{G}_k^{co} \equiv S_{k-1}^{-1} G_k^{co} S_{k-1}^{-T} = \check{S}_{k-1}^T G_k^{cr} \check{S}_{k-1} \\ &= \begin{bmatrix} \Sigma_k^{-1/2} V_k^T R_k^T E_{k-1}^T L_k^T L_k E_{k-1} R_k V_k \Sigma_k^{-1/2} & \Sigma_k^{-1/2} V_k^T R_k^T E_{k-1}^T L_k^T L_k A_{k-1} \widetilde{R}_{k-1} \widetilde{V}_{k-1} \Theta_{k-1}^{-1/2} \\ \Theta_{k-1}^{-1/2} \widetilde{V}_{k-1}^T \widetilde{R}_{k-1}^T A_{k-1}^T L_k^T L_k E_{k-1} R_k V_k \Sigma_k^{-1/2} & \Theta_{k-1}^{-1/2} \widetilde{V}_{k-1}^T \widetilde{R}_{k-1}^T A_{k-1}^T L_k^T L_k A_{k-1} \widetilde{R}_{k-1} \widetilde{V}_{k-1} \Theta_{k-1}^{-1/2} \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_k & 0 \\ 0 & 0 \end{bmatrix}.\end{aligned}$$

On the other hand, one can also show that the noncausal reachability and observability Gramians of the transformed periodic descriptor system (7.1) satisfy

$$\widehat{G}_k^{nr} \equiv T_k^{-1} G_k^{nr} T_k^{-T} = \begin{bmatrix} 0 & 0 \\ 0 & \Theta_k \end{bmatrix} = S_k^{-1} G_{k+1}^{no} S_k^{-T} \equiv \widehat{G}_{k+1}^{no}, \quad k = 0, 1, \ldots, K-1.$$

Consequently, S_k and T_k $(k = 0, 1, \ldots, K-1)$ are the desired balancing transformations such that the realization (7.1) is balanced. In summary, we have the following theorem.

Theorem 7.1. *For completely reachable and completely observable periodic discrete-time descriptor system (1.1) with $\{(E_k, A_k)\}_{k=0}^{K-1}$ being pd-stable, there exist nonsingular periodic matrices S_k and T_k $(k = 0, 1, \ldots, K-1)$ with $T_K \equiv T_0$ such that the transformed realization (7.1) is balanced.*

Remark. As in the cases of standard state space systems [58, 94] and descriptor systems [110, 113], the balancing transformation matrices for periodic descriptor system (1.1) are not unique. Indeed, if $\{(S_k, T_k)\}_{k=0}^{K-1}$ denotes a set of balancing transformation pairs

for the periodic descriptor system (1.1), then for any diagonal matrix D with diagonal entries ± 1, the set of matrix pairs $\{(S_k D, T_k D)\}_{k=0}^{K-1}$ are also the balancing transformation matrices for the periodic descriptor system (1.1).

8 Concluding Remarks

In this chapter we have derived the necessary and sufficient conditions for complete reachability and complete observability of periodic time-varying descriptor systems. Furthermore, the important concepts of reachability/observability Gramians, Hankel singular values and balanced realization have been generalized for periodic discrete-time descriptor systems. These are useful in the model reduction problem via the balanced truncation method.

In addition, in Theorem 4.1, the reachability/observability Gramians are shown to satisfy some projected GDPLE which can be computed numerically by applying the PQZ algorithm with reordering strategies. A numerical example is given to illustrate the feasibility and reliability of the proposed algorithm in Section 5.

NATIONAL TSING HUA UNIVERSITY

References

[1] J. Abels and P. Benner, *DAREX – a collection of benchmark examples for discrete-time algebraic Riccati equations (version 2.0)*, Tech. Rep. SLICOT Working Note 1999-16, The Working Group on Software, 1999.

[2] G. Ammar and V. Mehrmann, *On Hamiltonian and symplectic Hessenberg forms*, **Lin. Alg. Appl.**, 149 (1991), pp. 55–72.

[3] B. D. O. Anderson, *Second-order convergent algorithms for the steady-state Riccati equation*, **Int. J. Control**, 28 (1978), pp. 295–306.

[4] M. Athans, W. Levine, and A. Levis, *A system for the optimal and suboptimal position and velocity control for a string of high-speed vehicles*, in **Proc. 5th Int. Analogue Computation Meetings**, Lausanne, Switzerland, 1967.

[5] Z. Bai and J. Demmel, *On swapping diagonal blocks in real Schur form*, **Lin. Alg. Appl.**, 186 (1993), pp. 73–95.

[6] Z. Bai and J. Demmel, *Using the matrix sign function to compute invariant subspaces*, **SIAM J. Matrix Analy. Appl.**, 19 (1998), pp. 205–225.

[7] Z. Bai, J. Demmel, and M. Gu, *An inverse free parallel spectral divide and conquer algorithm for nonsymmetric eigenproblems*, **Numer. Math.**, 76 (1997), pp. 279–308.

[8] Z. Bai and Q. Qian, *Inverse free parallel method for the numerical solution of algebraic Riccati equations*, in **Proc. Fifth SIAM Conf. Appl. Lin. Alg., Snowbird, UT**, J. G. Lewis, ed., SIAM, Philadelphia, PA, 1994, pp. 167–171.

[9] L. Balzer, *Accelerated convergence of the matrix sign function*, **Int. J. Control**, 21 (1980), pp. 1057–1078.

[10] A. Y. Barraud, *Investigation autour de la fonction signe d'une matrice, application á l'équation de Riccati*, **R.A.I.R.O. Automatique**, 13 (1979), pp. 335–368.

[11] A. Y. Barraud, *Produit étoile et fonction signe de matrice. application á l'équation de Riccati dans le cas discret*, **R.A.I.R.O. Automatique**, 14 (1980), pp. 55–85.

[12] M. S. Bazaraa, H. D. Sheraii, and C. M. Shetty, **Nonlinear Programming**, John Wiley & Sons, 1993.

[13] A. N. Beavers and E. D. Denman, *Asymptotic solutions to the matrix Riccati equation*, **Mathematical Biosciences**, 20 (1974), pp. 339–344.

[14] A. N. Beavers and E. D. Denman, *A computational method for eigenvalues and eigenvectors of a matrix with real eigenvalues*, **Numer. Math.**, 21 (1974), pp. 389–396.

[15] A. N. Beavers and E. D. Denman, *A new similarity transformation method for eigenvalues and eigenvectors*, **Mathematical Biosciences**, 21 (1974), pp. 143–169.

[16] A. N. Beavers and E. D. Denman, *A new solution method for matrix quadratic equations*, **Mathematical Biosciences**, 20 (1974), pp. 135–143.

[17] D. J. Bender, *Lyapunov-like equations and reachability/observability gramians for descriptor systems*, **IEEE Trans. Auto. Control**, 32 (1987), pp. 343–348.

[18] P. Benner, *Contributions to the numerical solutions of algebraic Riccati equations and related eigenvalue problems*, PhD Dissertation, Fakultät für Mathematik, TU Chemnitz-Zwickau, Chemnitz, Germany, 1997.

[19] P. Benner and R. Byers, *Evaluating products of matrix pencils and collapsing matrix products*, **Num. Lin. Alg. Appl.**, 8 (2001), pp. 357–380.

[20] P. Benner, R. Byers, R. Mayo, E. S. Quintana-Orti, and V. Hernández, *Parallel algorithms for LQ optimal control of discrete-time periodic linear systems*, **J. Para. Distr. Comp.**, 62 (2002), pp. 306–325.

[21] P. Benner, A. J. Laub, and V. Mehrmann, *A collection of benchmark examples for the numerical solution of algebraic Riccati equations II: Discrete-time case*, Tech. Rep. SPC 95_23, Fakultät für Mathematik, TU Chemnitz–Zwickau, 09107 Chemnitz, FRG, 1995. Available from `http://www.tu-chemnitz.de/sfb393/spc95pr.html`.

[22] P. Benner, A. J. Laub, and V. Mehrmann, *A collection of benchmark examples for the numerical solution of algebraic Riccati equations I: Continuous-time case*, Tech. Rep. SPC 95_22, Fakultät für Mathematik, TU Chemnitz–Zwickau, 09107 Chemnitz, FRG, 1995. Available from `http://www.tu-chemnitz.de/sfb393/spc95pr.html`.

[23] P. Benner, R. Mayo, E. S. Quintana-Orti, and V. Hernández, *A coarse grain parallel solver for periodic Riccati equations*, Tech. Rep. 2000-01, Depto. de Informática, 12080-Castellón, Spain, 2000.

[24] P. Benner, R. Mayo, E. S. Quintana-Orti, and V. Hernández, *Solving diacrete-time periodic Riccati equations on a cluster*, in **Euro-Par 2000 Parallel Processing**, A. Bode, T. Ludwig, W. Karl, and R. Wismuller, eds., no. 1900 in Lecture Notes in Computer Science, Springer-Verlag, 2000, pp. 824–828.

[25] P. Benner, V. Mehrmann, V. Sima, S. V. Huffel, and A. Varga, *SLICOT - a subroutine library in systems and control theory*, **Applied and Computational Control, Signals, and Circuits**, 1 (1999), pp. 499–539.

[26] M. C. Berg, N. Amit, and J. D. Powell, *Multirate digital control system design*, **IEEE Trans. Auto. Control**, 33 (1988), pp. 1139–1150.

[27] W. Bialkowski, *Application of steady-state Kalman filters — theory with field results*, in **Proc. Joint Automat. Cont. Conf.**, Philadelphia, PA, 1978.

[28] S. Bittanti, *Deterministic and stochastic linear periodic systems*, in **Time Series and Linear Systems**, S. Bittanti, ed., Springer Verlag, New York, 1986, pp. 141–182.

[29] S. Bittanti and P. Colaneri, *Analysis of discrete-time linear periodic systems*, in **Control and Dynamic Systems**, C. T. Leondes, ed., vol. 78, Academic Press, New York, 1996.

[30] S. Bittanti and P. Colaneri, *Periodic control*, in **Wiley Encyclopedia of Electrical and Electronic Engineering**, J. G. Webster, ed., vol. 16, Wiley, New York, 1999, pp. 59–74.

[31] S. Bittanti, P. Colaneri, and G. D. Nicolao, *The difference periodic Riccati equation for the periodic prediction problem*, **IEEE Trans. Auto. Control**, 33 (1988), pp. 706–712.

[32] S. Bittanti, P. Colaneri, and G. D. Nicolao, *The periodic Riccati equation*, in **The Riccati Equation**, S. Bittanti, A. Laub, and J. Willems, eds., Springer-Verlag, 1991, pp. 127–162.

[33] A. Bojanczyk, G. H. Golub, and P. Van Dooren, *The periodic Schur decomposition. Algorithms and applications*, in **Proc. SPIE Conference**, vol. 1770, San Diego, 1992, pp. 31–42.

[34] J. H. Brandts, *Matlab code for sorting real Schur forms*, **Num. Lin. Alg. Appl.**, 9 (2002), pp. 249–261.

[35] R. Bru, C. Coll, and N. Thome, *Compensating periodic descriptor systems*, **Sys. Contr. Lett.**, 43 (2001), pp. 133–139.

[36] A. Bunse-Gerstner, R. Byers, and V. Mehrmann, *A chart of numerical methods for structured eigenvalue problems*, **SIAM J. Matrix Analy. Appl.**, 13 (1992), pp. 419–453.

TSING HUA UNIVERSITY

[37] A. Bunse-Gerstner, V. Mehrmann, and N. K. Nichols, *Regularization of descriptor systems by derivative and proportional state feedback*, **SIAM J. Matrix Analy. Appl.**, 13 (1992), pp. 46–67.

[38] A. Bunse-Gerstner, V. Mehrmann, and Watkins, *An SR algorithm for Hamiltonian matrices, based on Gaussian elimination*, **Methods of Operations Research**, 58 (1989), pp. 15–26.

[39] R. Byers, *A Hamiltonian QR-algorithm*, **SIAM J. Sci. Statist. Comput.**, 7 (1986), pp. 212–229.

[40] R. Byers, *Numerical stability and instability in matrix sign function based algorithms*, in **Computational and Combinatorial Methods in System Theory**, C. Byrnes and A. Lindquist, eds., North-Holland, 1986, pp. 185–200.

[41] R. Byers, *Solving the algebraic Riccati equation with the matrix sign function*, **Lin. Alg. Appl.**, 85 (1987), pp. 267–279.

[42] R. Byers, C. He, and V. Mehrmann, *the matrix sign function method and the computation of invariant subspaces*, **SIAM J. Matrix Analy. Appl.**, 18 (1997), pp. 615–632.

[43] R. Byers and N. Rhee, *Cyclic Schur and Hessenberg Schur numerical methods for solving periodic Lyapunov and Sylvester equations*, Technical Report, Dept. of Mathematics, Univ. of Missouri at Kansas City, 1995.

[44] E. K.-W. Chu, H.-Y. Fan, and W.-W. Lin, *A generalized structure-preserving doubling algorithm for generalized discrete-time algebraic Riccati equations*, preprint 2002-29, NCTS, National Tsing Hua University, Hsinchu 300, Taiwan, 2003.

[45] E. K.-W. Chu, H.-Y. Fan, and W.-W. Lin, *A structure-preserving doubling algorithm for continuous-time algebraic Riccati equations*, preprint 2002-28, NCTS, National Tsing Hua University, Hsinchu 300, Taiwan, 2003.

[46] E. K.-W. Chu, H.-Y. Fan, W.-W. Lin, and C.-S. Wang, *A structure-preserving doubling algorithm for periodic descrete-time algebraic Riccati equations*, preprint 2002-18, NCTS, National Tsing Hua University, Hsinchu 300, Taiwan, 2003.

[47] R. E. Crochiere and L. R. Rabiner, **Mutirate Digital Signal Processing**, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[48] L. Dai, **Singualr Control Systems**, Springer-Verlag Berlin, Heidelberg, 1989.

[49] E. J. Davison and W. Gesing, *The systematic design of control systems for the multivariable servomechenism problem*, in **Alternatives for Linear Multivariable Control**, M. K. Sain and J. L. Peczkowsky, eds., Nat. Eng. Consortium Inc., Chicago, IL, 1978.

[50] E. Denman and R. Beavers, *The matrix sign function and computations in systems*, **Appl. Math. Comp.**, 2 (1976), pp. 63–94.

[51] L. Dieci, *Some numerical considerations and Newton's method revisited for solving algebraic Riccati equations*, **IEEE Trans. Auto. Control**, 36 (1991), pp. 608–616.

[52] A. Feuer and G. C. Goodwin, **Sampling in Digital Signal Processing and Control**, Birkhauser, New York, 1996.

[53] D. S. Flamm and A. J. Laub, *A new shift-invariant representation of perioidc linear systems*, **Sys. Contr. Lett.**, 17 (1991), pp. 9–14.

[54] C. Foulard, S. Gentil, and J. P. Sandraz, **Commande et régulation par calculateur numérique: De la théorie aux applications**, Eyrolles, Paris, 1977.

[55] B. Francis and T. T. Georgiou, *Stability theory for linear time-invariant plants with periodic digital controllers*, **IEEE Trans. Auto. Control**, 33 (1988), pp. 820–832.

[56] J. Gardiner and A. J. Laub, *A generalization of the matrix-sign-function solution to the algebraic Riccati equations*, **Int. J. Control**, 44 (1986), pp. 823–832.

[57] W. A. Gardner, **Cyclostationarity in Communnications and Signal Processing**, IEEE Press, New York, 1994.

[58] K. Glover, *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-errors bounds*, **Int. J. Control**, 39 (1984), pp. 1115–1193.

[59] G. H. Golub and C. F. Van Loan, **Matrix Computations, 3rd ed.**, The Johns Hopkins University Press, 1996.

[60] T. Gudmundsson, C. Kenney, and A. J. Laub, *Scaling of the discrete-time algebraic Riccati equation to enhance stability of the Schur solution method*, **IEEE Trans. Auto. Control**, 37 (1992), pp. 513–518.

[61] S. Hammarling, *Newton's method for solving the algebraic Riccati equation*, NPL Rep. DITC 12/82, Nat. Phys. Lab., Teddington, Middlesex TW11 0LW, U.K., 1982.

[62] J. J. Hench and A. J. Laub, *Numerical solution of the discrete-time periodic Riccati equation*, **IEEE Trans. Auto. Control**, 39 (1994), pp. 1197–1210.

[63] N. J. Higham, **Accuracy and Stability of Numerical Algorithms**, SIAM, Philadelphia, PA, 1996.

[64] J. L. Howland, *The sign matrix and the separation of matrix eigenvalues*, **Lin. Alg. Appl.**, 49 (1983), pp. 221–332.

[65] P. Hr. Petkov, N. D. Christov, and M. M. Konstantinov, *On the numerical properties of the Schur approach for solving the matrix Riccati equation*, **Sys. Contr. Lett.**, 9 (1987), pp. 197–201.

[66] G. D. Ianculescu, J. Ly, A. J. Laub, and P. M. Papadopoulos, *Space station freedom solar array H_∞ control.* Talk at 31st IEEE Conf. on Decision and Control, Tucson, AZ, Dec. 1992.

[67] R. W. Isniewski and M. Blanke, *Fully magnetic attitude control for spacecraft subject to gravity gradient*, **Automatica**, 35 (1999), pp. 1201–1214.

[68] W. Johnson, **Helicopter Theory**, Princeton University Press, Princeton, NJ, 1996.

[69] M. Kimura, *Convergence of the doubling algorithm for the discrete-time algebraic Riccati equation*, **Int. J. Syst. Sci.**, 19 (1988), pp. 701–711.

[70] M. Kimura, *Doubling algorithm for continuous-time algebraic Riccati equation*, **Int. J. Syst. Sci.**, 20 (1989), pp. 191–202.

[71] D. Kleinman, *On an iterative technique for Riccati equation computations*, **IEEE Trans. Auto. Control**, AC-13 (1968), pp. 114–115.

[72] M. Kono, *Eigenvalue assignment in linear discrete-time system*, **Int. J. Control**, 32 (1980), pp. 149–158.

[73] Y.-C. Kuo, W.-W. Lin, and S.-F. Xu, *Regularrization of linear discrete-time perioidc descriptor systems by derivative and proportional state feedback*. To appear in **SIAM J. Matrix Analy. Appl.**, 2004.

[74] D. Lainiotis, N. Assimakis, and S. Katsikas, *New doubling algorithm for the discrete periodic Riccati equation*, **Appl. Maths. Comp.**, 60 (1994), pp. 265–283.

[75] A. J. Laub, *A Schur method for solving algebraic Riccati equations*, **IEEE Trans. Auto. Control**, 24 (1979), pp. 913–921.

[76] A. J. Laub, *Algebraic aspects of generalized eigenvalue problems for solving Riccati equations*, in **Computational and Combinatorial Methods in Systems Theory**, C. I. Byrnes and A. Lindquist, eds., Elsevier (North-Holland), 1986, pp. 213–227.

[77] A. J. Laub, *Invariant subspace methods for the numerical solution of Riccati equations*, in **The Riccati Equation**, S. Bittanti, A. J. Laub, and J. C. Willems, eds., Springer-Verlag, Berlin, 1991, pp. 163–196.

[78] A. J. Laub, M. T. Heath, C. C. Paige, and R. C. Ward, *Computation of system balancing transformations and other applications of simutaneous diagonalization algorithms*, **IEEE Trans. Auto. Control**, 32 (1987), pp. 115–122.

[79] F. L. Lewis, *Fundamental, reachability, and observability matrices for discrete descriptor systems*, **IEEE Trans. Auto. Control**, 30 (1985), pp. 502–505.

[80] W.-W. Lin and J.-G. Sun, *Perturbation analysis for eigenproblem of periodic matrix pairs*, **Lin. Alg. Appl.**, 337 (2001), pp. 157–187.

[81] W.-W. Lin and J.-G. Sun, *Perturbation analysis of the periodic discrete-time algebraic Riccati equation*, **SIAM J. Matrix Analy. Appl.**, 24 (2002), pp. 411–438.

[82] W.-W. Lin, C.-S. Wang, and Q.-F. Xu, *Numerical computation of the minimum H_∞ norm of the discrete-time output feedback control problem*, **SIAM J. Numer. Anal.**, 38 (2000), pp. 515–547.

[83] M.-L. Liou and Y.-L. Kuo, *Exact analysis of switched capacitor circuits with arbitrary inputs*, **IEEE Trans. Circuits and Systems**, 26 (1979), pp. 213–223.

[84] L.-Z. Lu and W.-W. Lin, *An iterative algorithm for the solution of the discrete time algebraic Riccati equations*, **Lin. Alg. Appl.**, 189 (1993), pp. 465–488.

[85] L.-Z. Lu, W.-W. Lin, and C. E. M. Pearce, *An efficient algorithm for the discrete-time algebraic Riccati equation*, **IEEE Trans. Auto. Control**, 44 (1999), pp. 1216–1220.

[86] A. N. Malyshev, *Parallel algorithm for solving some spectral problems of linear algebra*, **Lin. Alg. Appl.**, 188/189 (1993), pp. 489–520.

[87] A. Marzollo, **Periodic Optimization**, Springer Verlag, Berlin, 1972.

[88] MathWorks, **MATLAB user's guide (for UNIX Workstations)**, The Math Works, Inc., 1992.

[89] R. McKillip, *Periodic model following controller for the control-configured helicopter*, **Journal of the American Helicopter Society**, 36 (1991), pp. 4–12.

[90] V. Mehrmann, *A symplectic orthogonal method for single input or single output discrete time optimal linear quadratic control problems*, **SIAM J. Matrix Analy. Appl.**, (1988), pp. 221–248.

[91] V. Mehrmann, **The Autonomous Linear Quadratic Control Problem**, Springer-Verlag, 1991.

[92] V. Mehrmann, *A step toward a unified treatment of continuous and discrete time control problems*, **Lin. Alg. Appl.**, 241-243 (1996), pp. 749–779.

[93] V. Mehrmann and E. Tan, *Defect correction methods for the solution of algebraic Riccati equations*, **IEEE Trans. Auto. Control**, AC-33 (1988), pp. 695–698.

[94] B. C. Moore, *Principal component analysis in linear systems: controllability, observability, and model reduction*, **IEEE Trans. Auto. Control**, 26 (1981), pp. 17–32.

[95] C. C. Paige and C. F. Van Loan, *A Schur decomposition for Hamiltonian matrices*, **Lin. Alg. Appl.**, 41 (1981), pp. 11–32.

[96] T. Pappas, A. J. Laub, and N. R. Sandell, *On the numerical solution of the discrete-time algebraic Riccati equation*, **IEEE Trans. Auto. Control**, 25 (1980), pp. 631–641.

[97] L. Patnaik, N. Viswanadham, and I. Sarma, *Computer control algorithms for a tubular ammonia reactor*, **IEEE Trans. Auto. Control**, 25 (1980), pp. 642–651.

[98] T. Penzl, *Numerical solution of generalized Lyapunov equations*, **Adv. Comput. Math.**, 8 (1998), pp. 33–48.

[99] P. Petkov, N. Christov, and M. Konstantinov, *A posteriori error analysis of the generalized Schur approach for solving the discrete matrix Riccati equation*, preprint,

Department of Automatics, Higher Institute of Mechenical and Electrical Engineering, 1756 Sofia, Bulgaria, 1989.

[100] M. E. Pittelkau, *Optimal periodic control for spacecraft pointing and attitude determination*, **J. of Guidance, Control, and Dynamics**, 16 (1993), pp. 1078–1084.

[101] J. A. Richards, **Analysis of Periodically Time-Varying Systems**, Springer-Verlag, Berlin, 1983.

[102] J. Roberts, *Linear model reduction and solution of the algebraic Riccati equation by the use of the sign function*, **Int. J. Control**, 32 (1980), pp. 667–687.

[103] N. Sandell, *On Newton's method for Riccati equation solution*, **IEEE Trans. Auto. Control**, AC-19 (1974), pp. 254–255.

[104] V. Sima, **Algorithms for Linear-Quadratic Optimization, volume 200 of Pure and Applied Mathematics**, Marcel Dekker, Inc., New York, NY, 1996.

[105] J. Sreedhar and P. Van Dooren, *Periodic Schur form and some matrix equations*, **Systems and Networks: Mathematical Theory and Applications**, 77 (1994), pp. 339–362.

[106] J. Sreedhar and P. Van Dooren, *Forward/backward decomposition of periodic descriptor systems and two point boundary value problems*, in **European Control Conf.**, 1997.

[107] J. Sreedhar and P. Van Dooren, *Periodic descriptor systems: solvability and conditionability*, **IEEE Trans. Auto. Control**, 44 (1999), pp. 310–313.

[108] G. W. Stewart, *HQR3 and EXCHNG: Fortran subroutines for calculating and ordering the eigenvalues of a real upper Hessenberg matrix*, **ACM Trans. Math. Software**, 2 (1976), pp. 275–280.

[109] G. W. Stewart and J.-G. Sun, **Matrix Perturbation Theory**, Academic Press, New York, 1990.

[110] T. Stykel, *Model reduction of descriptor systems*, Technical Report 720-2001, Institut für Mathematik, TU Berlin, D-10263 Berlin, Germany, 2001.

[111] T. Stykel, *Analysis and numerical solution of generalized Lyapunov equations*, PhD Dissertation, Institut für Mathematik, Tecnische Universität Berlin, Berlin, 2002.

[112] T. Stykel, *Stability and inertia theorems for generalized Lyapunov equations*, **Lin. Alg. Appl.**, 355 (2002), pp. 297–314.

[113] T. Stykel, *Balanced truncation model reduction for semidiscretized Stokes equation*, Technical Report 04-2003, Institut für Mathematik, TU Berlin, D-10263 Berlin, Germany, 2003.

[114] T. Stykel, *Input-output invariants for descriptor systems*, preprint PIMS-03-1, Pacific Institute for the Mathematical Sciences, Canada, 2003.

[115] J.-G. Sun, *Sensitivity analysis of the discrete-time algebraic Riccati equation*, **Lin. Alg. Appl.**, 275/276 (1998), pp. 595–615.

[116] L. Tong, G. Xu, and T. Kailath, *Blind identification and equalization based on second-order statistics: A time domain approach*, **IEEE Trans. Information Theory**, 40 (1994), pp. 340–349.

[117] P. P. Vaidyanathan, *Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial*, in **Proc. IEEE**, vol. 78, 1990, pp. 56–93.

[118] P. P. Vaidyanathan, **Mutirate Systems and Filter-Banks**, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[119] P. Van Dooren, *A generalized eigenvalue approach for solving Riccati equations*, **SIAM J. Sci. Statist. Comput.**, 2 (1981), pp. 121–135.

[120] P. Van Dooren, *Two point boundary value and periodic eigenvalue problems*, in **Proc. 1999 IEEE Intel. Symp. CACSD, Kohala Coast-Island, Hawaii, USA**, K. Kirchgassner et al., ed., August 1999, pp. 22–27.

[121] P. Van Dooren and J. Sreedhar, *When is a periodic discrete-time system equivalent to a time invariant one?*, **Lin. Alg. Appl.**, 212/213 (1994), pp. 131–151.

[122] A. Varga, *Periodic Lyapunov equations: some applications and new algorithms*, **Int. J. Control**, 67 (1997), pp. 69–87.

[123] A. Varga, *Balancing related methods for minimal realization of perioidc systems*, **Sys. Contr. Lett.**, 36 (1999), pp. 339–349.

[124] A. Varga, *Balanced truncation model reduction of periodic systems*, in **Proc. of IEEE Conference on Decision and Control, Sydney, Australia**, 2000.

[125] A. Varga, *Robust and minimum norm pole assignment with periodic state feedback*, **IEEE Trans. Auto. Control**, 45 (2000), pp. 1017–1022.

[126] A. Varga and S. Pieters, *Gradient-based approach to solve optimal periodic output feedback control problems*, **Automatica**, 34 (1998), pp. 477–481.

[127] A. Varga and P. Van Dooren, *Computing the zeros of periodic descriptor systems*, **Sys. Contr. Lett.**, 50 (2003), pp. 371–381.

[128] J. Vlach, K. Singhai, and M. Vlach, *Computer oriented formulation of equations and analysis of switched-capacitor networks*, **IEEE Trans. Circuits and Systems**, 31 (1984), pp. 735–765.

[129] J. Xin, H. Kagiwada, A. Sano, H. Tsuj, and S. Yoshimoto, *Regularization approach for detection of cyclostationary signals in antenna array processing*, in **IFAC Symposium on System Identification**, vol. 2, 1997, pp. 529–534.

[130] V. A. Yakubovich and V. M. Starzhinskii, **Linear Differential Equations with Perioidc Coefficients**, Wiley, New York, 1975.

[131] K. Zhou, J. C. Doyle, and K. Glover, **Robust and Optimal Control**, Prentice-Hall, Upper Saddle River, NJ, 1996.