

# Schur Complement Based Preconditioners for Twofold and Block Tridiagonal Saddle Point Problems

Mingchao Cai<sup>1,\*</sup>, Guoliang Ju<sup>2</sup> and Jingzhi Li<sup>3</sup>

<sup>1</sup> Department of Mathematics, Morgan State University, Baltimore, MD 21251, USA.

<sup>2</sup> Solid State R&D Department, Shenzhen Tenfong Co. Ltd., Shenzhen 518055, P.R. China.

<sup>3</sup> Department of Mathematics, Southern University of Science and Technology, Shenzhen 518055, P.R. China.

Received 30 December 2023; Accepted 7 April 2024

---

**Abstract.** In this paper, we consider using Schur complements to design preconditioners for twofold and block tridiagonal saddle point problems. One type of the preconditioners are based on the nested (or recursive) Schur complement, the other is based on an additive type Schur complement after permuting the original saddle point systems. We analyze different preconditioners incorporating the exact Schur complements. We show that some of them will lead to positively stable preconditioned systems if proper signs are selected in front of the Schur complements. These positive-stable preconditioners outperform other preconditioners if the Schur complements are further approximated inexactly. Numerical experiments for a 3-field formulation of the Biot model are provided to verify our predictions.

**AMS subject classifications:** 65F08, 65M22, 74F20

**Key words:** Schur complement, block tridiagonal systems, positively stable preconditioners, Routh-Hurwitz stability criterion.

---

\*Corresponding author. *Email addresses:* cmchao2005@gmail.com (M. Cai), jugl@tenfong.cn (G. Ju), li.jz@sustech.edu.cn (J. Li)

## 1 Introduction

Many application problems will lead to twofold and/or block tridiagonal saddle point linear systems. Important examples include mixed formulations of the Biot model [1, 9, 22, 26, 33], the coupling of fluid flow with porous media flow [10, 21, 29], hybrid discontinuous Galerkin approximation of Stokes problem [17], liquid crystal problem [3, 36] and optimization problems [23, 26, 28, 34, 38]. Some of these problems (or after permutations) will lead to a twofold saddle point problem [3, 12, 21, 23, 24, 38, 40, 41] (or the so-called double saddle point problem) of the following form:

$$\mathcal{A}x = \begin{bmatrix} A_1 & B_1^T & 0 \\ C_1 & -A_2 & B_2^T \\ 0 & C_2 & A_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}. \quad (1.1)$$

A negative sign in front of  $A_2$  is just for the ease of notation. After simple permutations, the system matrix of (1.1) can be rewritten into the following form:

$$\mathcal{A} = \begin{bmatrix} A_1 & 0 & B_1^T \\ 0 & A_3 & C_2 \\ C_1 & B_2^T & -A_2 \end{bmatrix}. \quad (1.2)$$

We call the system matrix in (1.2) permutation-equivalent to that in (1.1). Without causing confusion, we continue to use the notation  $\mathcal{A}$  for the permuted matrix (1.2). The linear system in (1.2) arises naturally from the domain decomposition methods [30, 39]. In this work, we only assume that  $A_1$  is invertible and the global system matrix  $\mathcal{A}$  is invertible. Many special cases, e.g. if  $A_2 = 0$ , or  $A_3 = 0$ , or  $A_2 = A_3 = 0$  can be cast into the above forms of twofold saddle point systems. Our discussions will try to cover all these special cases.

The above 3-by-3 block linear problems (1.1) and (1.2) can be naturally extended to  $n$ -tuple cases. For example, when the system matrix in (1.1) is extended to the  $n$ -tuple case, it is the block tridiagonal systems discussed in [38]. When the system matrix in (1.2) is extended to the  $n$ -tuple case, it corresponds to the linear system resulting from the domain decomposition method for elliptic problems with  $n-1$  subdomains. In many references, these linear systems are assumed to be symmetric. No matter whether it is symmetric or not,  $\mathcal{A}$  is generally indefinite. For solving such a system in large-scale computations, Krylov subspace methods with preconditioners are usually applied. The analysis in [25, 31] indicates that one should employ Schur complement based preconditioners [5, 15, 16, 27] and

then resort to replacing each exact Schur complement with an approximate one. If a Krylov subspace method satisfies an optimality or Galerkin property (such as conjugate gradient method and minimal residual methods [18, 31, 37]), the number of iterations depends on the degree of the minimal polynomial that the pre-conditioned matrix  $\mathcal{T}$  satisfies [25, 31, 37]. More precisely, the number of iterations depends on the dimension of the Krylov subspace  $\text{Span}\{\mathbf{r}, \mathcal{T}\mathbf{r}, \mathcal{T}^2\mathbf{r}, \mathcal{T}^3\mathbf{r}, \dots\}$ . For a classical 2-by-2 saddle point problem, by employing block-diagonal or block-triangular preconditioners with the (2,2) block being the Schur complement, it has been shown that the corresponding preconditioned system satisfies a polynomial of the degree of 4 or 2 [25, 31]. Therefore, devising preconditioners based on Schur complements is of vital importance.

In this study, we explore two methodologies for designing preconditioners tailored for 3-by-3 block systems exhibiting block tridiagonal form, as well as their extensions to  $n$ -tuple scenarios. One approach centers around the nested (or recursive) Schur complement [38]. Unlike the approach presented in [38], our work adopts a more expansive assumption, allowing  $A_1, A_2$ , and  $A_3$  to be non-symmetric or negative definite. Another approach is founded on the permuted matrix (1.2). Initially treating the first 2-by-2 block as a unified entity, albeit potentially indefinite, we transform the system into the saddle point form discussed in [25]. Subsequently, we explore preconditioners based on an additive-type Schur complement [14, 30, 39]. For 3-by-3 block systems, the additive-type Schur complement-based preconditioner yields a system that satisfies a polynomial of degree 2 (block triangular) or 4 (block diagonal), while preconditioners based on the nested Schur complement satisfy a polynomial with a degree of 3 (block triangular) or 6 (block diagonal). When extended to  $n$ -tuple scenarios, if the preconditioned system is diagonalizable, the degree of the polynomial satisfied by a nested Schur complement-based preconditioner is significantly higher than that satisfied by an additive Schur complement-based preconditioner. Thus, from an algebraic perspective, when employing the GMRES method, the additive-type preconditioner is favored. The generalization of the form (1.2) to  $n$ -tuple cases corresponds to the linear system resulting from domain decomposition methods, referencing some literature results on domain decomposition methods.

Commencing with a twofold saddle point problem, we generalize our theory to  $n$ -tuple block tridiagonal saddle point problems. Our study demonstrates that judiciously selecting signs in front of Schur complements in preconditioners results in a positively stable preconditioned system [16]. By using the Routh-Hurwitz stability criterion, we prove that the appropriate sign in front of each Schur complement yields a positively stable preconditioned system [16]. For positively stable preconditioners, we highlight that they outperform other precondi-

tioners if further inexact approximations are applied to the corresponding Schur complements. Numerical experiments, exemplified by a 3-field formulation of the Biot model, underscore the superior performance of positively stable preconditioners compared to other alternatives. Notably, we refrain from assuming the positive definiteness of each  $A_i$ , and the linear systems investigated in this study extend beyond saddle point problems. In the concluding remarks, we discuss and highlight the potential advantages and disadvantages of the two types of preconditioners, particularly when considering their corresponding inexact versions.

The outline of the remainder of this paper is as follows. In Section 2, we briefly recall the classic saddle point problem and its Schur complement, and introduce the twofold saddle point problem and the form of Schur complement, we then construct and analyze the block-triangular and block-diagonal preconditioners based on Schur complement for twofold saddle point problems. Furthermore, we extend these results to the  $n$ -tuple saddle point problem in Section 3. Some additive Schur complement based preconditioners are constructed and the corresponding known results in the literature are recalled in Section 4 for twofold saddle point problems. Generalizations to  $n$ -tuple cases are provided in Section 5. In Section 6, numerical experiments for a 3-field formulation of the Biot model are provided to justify the advantages of using positively stable preconditioners. Finally, concluding remarks are given in Section 7.

## 2 Nested Schur complement based preconditioners for twofold saddle point systems

The matrix form of a classic saddle point problem [5, 25] reads as

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ C & -D \end{bmatrix}.$$

Following the notation in [38], the Schur complement of  $\mathcal{A}$  is defined as

$$\text{Schur}(\mathcal{A}) = -D - CA^{-1}B^T.$$

In many references,  $A$  is assumed to be a symmetric positive definite, if  $C$  is assumed to be  $B$  and  $D$  is zero, then the system is called KKT system or saddle point system. Usually,  $D$  is assumed to be a symmetric and semi-positive definite. In this paper, we only make some assumptions that can guarantee the invertibility of  $\mathcal{A}$ . We assume that  $A$  and the Schur complement  $\text{Schur}(\mathcal{A})$  are invertible. For the twofold saddle point problem (1.1), following the notations in [38], we denote

$$S_1 = A_1, \quad S_2 = A_2 + C_1 S_1^{-1} B_1^T,$$

and denote the nested Schur complement as

$$S_3 = A_3 + C_2 S_2^{-1} B_2^T.$$

In [38], each  $S_i$  is assumed to be symmetric positive definite (SPD). From the expressions, if  $\mathcal{A}$  is symmetric and all  $A_i$ ,  $i=1,\dots,3$  are symmetric positive definite, then  $S_i$  are SPD. In this work, we only assume that each  $S_i$  is invertible. Such assumptions are sufficient enough to guarantee the invertibility of the global system. Under such an assumption, the block  $LDU$  decomposition of the system matrix is

$$\mathcal{A} = \begin{bmatrix} I & 0 & 0 \\ C_1 A_1^{-1} & I & 0 \\ 0 & -C_2 S_2^{-1} & I \end{bmatrix} \begin{bmatrix} A_1 & 0 & 0 \\ 0 & -S_2 & 0 \\ 0 & 0 & S_3 \end{bmatrix} \begin{bmatrix} I & A_1^{-1} B_1^T & 0 \\ 0 & I & -S_2^{-1} B_2^T \\ 0 & 0 & I \end{bmatrix}. \quad (2.1)$$

For ease of presentation, we will denote the above 3 block matrices as  $\mathcal{L}$ ,  $\mathcal{D}$ , and  $\mathcal{U}$ , respectively. We study both block-triangular and block-diagonal preconditioners for the system matrix (1.1). For block-triangular preconditioners, we focus on a lower triangular type with left preconditioning because an upper triangular one with right preconditioning can be discussed in a similar way [5, 25]. We consider the following preconditioner:

$$\mathcal{P}_{T_1} = \begin{bmatrix} A_1 & 0 & 0 \\ C_1 & -S_2 & 0 \\ 0 & C_2 & S_3 \end{bmatrix} = \mathcal{L}\mathcal{D}. \quad (2.2)$$

**Theorem 2.1.** *Assume that all  $S_i$ ,  $i=1,2,3$ , are invertible, then, for the preconditioner (2.2),*

$$\mathcal{P}_{T_1}^{-1} \mathcal{A} = \mathcal{U}, \quad (2.3)$$

*and  $\mathcal{P}_{T_1}^{-1} \mathcal{A}$  satisfies the polynomial equation  $(\lambda - 1)^3 = 0$ . If we further assume that  $B_1^T S_2^{-1} B_2^T$  is nonzero, then  $(\lambda - 1)^3$  is the minimal polynomial.*

*Proof.* From (2.1), because  $\mathcal{A} = \mathcal{L}\mathcal{D}\mathcal{U}$ ,  $\mathcal{P}_{T_1}^{-1} \mathcal{A} = (\mathcal{L}\mathcal{D})^{-1} \mathcal{A} = \mathcal{U}$ , which is the block upper triangular matrix of the form in (2.3). On one hand,  $(\mathcal{P}_{T_1}^{-1} \mathcal{A} - I)^3 = 0$ . To verify whether  $(\lambda - 1)^3$  is the minimal polynomial, we assume that there exist nonzero numbers  $a, b, c$  such that  $a\mathcal{U}^2 + b\mathcal{U} + cI = 0$ . We note that

$$\mathcal{U}^2 = \begin{bmatrix} I & 2A_1^{-1}B_1^T & -A_1^{-1}B_1^T S_2^{-1} B_2^T \\ 0 & I & -2S_2^{-1} B_2^T \\ 0 & 0 & I \end{bmatrix},$$

then

$$a\mathcal{U}^2 + b\mathcal{U} + cI = \begin{bmatrix} (a+b+c)I & (2a+b)A_1^{-1}B_1^T & -aA_1^{-1}B_1^T S_2^{-1}B_2^T \\ 0 & (a+b+c)I & -(2a+b)S_2^{-1}B_2^T \\ 0 & 0 & (a+b+c)I \end{bmatrix} = 0.$$

Because both  $A_1$  is invertible and  $B_1^T S_2^{-1}B_2^T$  is nonzero, from the (1,3) entry of  $a\mathcal{U}^2 + b\mathcal{U} + cI$ , we derive that  $a=0$ . Then, from the fact that the (1,2) and (2,3) entries of  $a\mathcal{U}^2 + b\mathcal{U} + cI$  are zeros, we conclude  $b=0$ . Lastly, from the diagonal entries, see that  $c=0$ . Therefore,  $(\lambda-1)^3$  is the minimal polynomial of  $\mathcal{P}_{T_1}^{-1}\mathcal{A}$ .  $\square$

Other preconditioners which will make the corresponding preconditioned systems have eigenvalues 1 or  $-1$  are

$$\begin{aligned} \mathcal{P}_{T_2} &= \begin{bmatrix} A_1 & 0 & 0 \\ C_1 & S_2 & 0 \\ 0 & -C_2 & S_3 \end{bmatrix}, \quad \mathcal{P}_{T_3} = \begin{bmatrix} A_1 & 0 & 0 \\ C_1 & S_2 & 0 \\ 0 & -C_2 & -S_3 \end{bmatrix}, \\ \mathcal{P}_{T_4} &= \begin{bmatrix} A_2 & 0 & 0 \\ C_1 & -S_2 & 0 \\ 0 & C_2 & -S_3 \end{bmatrix}. \end{aligned} \tag{2.4}$$

**Remark 2.1.** For preconditioners listed in (2.4), we have the following conclusions:

$$\mathcal{P}_{T_2}^{-1}\mathcal{A} = \begin{bmatrix} I & A_1^{-1}B_1^T & 0 \\ 0 & -I & S_2^{-1}B_2^T \\ 0 & 0 & I \end{bmatrix},$$

and  $\mathcal{P}_{T_2}^{-1}\mathcal{A}$  satisfies the polynomial equation  $(\lambda-1)^2(\lambda+1)=0$ ,

$$\mathcal{P}_{T_3}^{-1}\mathcal{A} = \begin{bmatrix} I & A_1^{-1}B_1^T & 0 \\ 0 & -I & S_2^{-1}B_2^T \\ 0 & 0 & -I \end{bmatrix},$$

and  $\mathcal{P}_{T_3}^{-1}\mathcal{A}$  satisfies the polynomial equation  $(\lambda-1)(\lambda+1)^2=0$ ,

$$\mathcal{P}_{T_4}^{-1}\mathcal{A} = \begin{bmatrix} I & A_1^{-1}B_1^T & 0 \\ 0 & I & -S_2^{-1}B_2^T \\ 0 & 0 & -I \end{bmatrix},$$

and  $\mathcal{P}_{T_4}^{-1}\mathcal{A}$  satisfies the polynomial equation  $(\lambda-1)^2(\lambda+1)=0$ .

**Definition 2.1.** A matrix is said to be stable if the real parts of all the eigenvalues are negative. Correspondingly, if the real parts of all the eigenvalues of a matrix  $A$  are positive, we say that these eigenvalues are positive real. If a preconditioner  $\mathcal{P}$  of  $\mathcal{A}$  satisfies  $\mathcal{P}^{-1}\mathcal{A}$  is positive real, we call  $\mathcal{P}$  is a positive stable preconditioner of  $\mathcal{A}$ .

A sufficient condition to guarantee the eigenvalues of a matrix  $A$  is positive real is the Hermitian part of  $A$ , i.e.,  $(A+A^H)/2$ , is SPD. The reverse direction is not true. There are many matrices whose eigenvalues are positive but their Hermitian parts are not SPD.

**Theorem 2.2.** Let the assumptions in Theorem 2.1 hold. We further assume that  $A_2=0$  and  $A_3=0$ . If the block-diagonal preconditioner is

$$\mathcal{P}_{D_1} = \begin{bmatrix} A_1 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & S_3 \end{bmatrix}, \quad (2.5)$$

then

$$\mathcal{T}_1 = \mathcal{P}_{D_1}^{-1} \mathcal{A} = \begin{bmatrix} I & A_1^{-1} B_1^T & 0 \\ S_2^{-1} C_1 & 0 & S_2^{-1} B_2^T \\ 0 & S_3^{-1} C_2 & 0 \end{bmatrix}, \quad (2.6)$$

which satisfies the polynomial equation  $(\lambda-1)(\lambda^2-\lambda-1)(\lambda^3-\lambda^2-2\lambda+1)=0$ .

*Proof.* By a simple calculation, one can derive that (2.6) holds. Then, we see that

$$\begin{aligned} \mathcal{T}_1^2 &= \begin{bmatrix} I + A_1^{-1} B_1^T S_2^{-1} C_1 & A_1^{-1} B_1^T & A_1^{-1} B_1^T S_2^{-1} B_2^T \\ S_2^{-1} C_1 & I + S_2^{-1} B_2^T S_3^{-1} C_2 & 0 \\ S_3^{-1} C_2 S_2^{-1} C_1 & 0 & I \end{bmatrix}, \\ \mathcal{T}_1^3 &= \begin{bmatrix} I + 2A_1^{-1} B_1^T S_2^{-1} C_1 & A_1^{-1} B_1^T (2I + S_2^{-1} B_2^T S_3^{-1} C_2) & A_1^{-1} B_1^T S_2^{-1} B_2^T \\ (2I + S_2^{-1} B_2^T S_3^{-1} C_2) S_2^{-1} C_1 & I & 2S_2^{-1} B_2^T \\ S_3^{-1} C_2 S_2^{-1} C_1 & 2S_3^{-1} C_2 & 0 \end{bmatrix}, \end{aligned}$$

and

$$\mathcal{T}_1 - I = \begin{bmatrix} 0 & A_1^{-1} B_1^T & 0 \\ S_2^{-1} C_1 & -I & S_2^{-1} B_2^T \\ 0 & S_3^{-1} C_2 & -I \end{bmatrix},$$

$$\begin{aligned}\mathcal{T}_1^2 - \mathcal{T}_1 - I &= \begin{bmatrix} -I + A_1^{-1} B_1^T S_2^{-1} C_1 & 0 & A_1^{-1} B_1^T S_2^{-1} B_2^T \\ 0 & S_2^{-1} B_2^T S_3^{-1} C_2 & -S_2^{-1} B_2^T \\ S_3^{-1} C_2 S_2^{-1} C_1 & -S_3^{-1} C_2 & 0 \end{bmatrix}, \\ \mathcal{T}_1^3 - \mathcal{T}_1^2 - 2\mathcal{T}_1 + I &= \begin{bmatrix} -I + A_1^{-1} B_1^T S_2^{-1} C_1 & A_1^{-1} B_1^T (-I + S_2^{-1} B_2^T S_3^{-1} C_2) & 0 \\ (-I + S_2^{-1} B_2^T S_3^{-1} C_2) S_2^{-1} C_1 & I - S_2^{-1} B_2^T S_3^{-1} C_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.\end{aligned}$$

Combining all of the above, one can verify that

$$\begin{aligned}(\mathcal{T}_1^2 - \mathcal{T}_1 - I)(\mathcal{T}_1 - I) &= (\mathcal{T}_1 - I)(\mathcal{T}_1^2 - \mathcal{T}_1 - I) \\ &= \begin{bmatrix} 0 & A_1^{-1} B_1^T S_2^{-1} B_2^T S_3^{-1} C_2 & -A_1^{-1} B_1^T S_2^{-1} B_2^T \\ S_2^{-1} B_2^T S_3^{-1} C_2 S_2^{-1} C_1 & -2S_2^{-1} B_2^T S_3^{-1} C_2 & 2S_2^{-1} B_2^T \\ -S_3^{-1} C_2 S_2^{-1} C_1 & 2S_3^{-1} C_2 & -I \end{bmatrix}.\end{aligned}$$

The  $(i,j)$  block of  $(\mathcal{T}_1 - I)(\mathcal{T}_1^2 - \mathcal{T}_1 - I)(\mathcal{T}_1^3 - \mathcal{T}_1^2 - 2\mathcal{T}_1 + I)$  is denoted as  $A_{ij}$ . Because  $A_2$  and  $A_3$  are zeros, and hence  $S_i = C_{i-1} S_{i-1}^{-1} B_{i-1}^T$ ,  $i=2,3$ . One can verify that

$$\begin{aligned}A_{11} &= A_1^{-1} B_1^T S_2^{-1} B_2^T S_3^{-1} C_2 (-I + S_2^{-1} B_2^T S_3^{-1} C_2) S_2^{-1} C_1 = 0, \\ A_{12} &= A_1^{-1} B_1^T S_2^{-1} B_2^T S_3^{-1} C_2 (I - S_2^{-1} B_2^T S_3^{-1} C_2) = 0, \\ A_{21} &= S_2^{-1} B_2^T S_3^{-1} C_2 S_2^{-1} C_1 (-I + A_1^{-1} B_1^T S_2^{-1} C_1) \\ &\quad - 2S_2^{-1} B_2^T S_3^{-1} C_2 (-I + S_2^{-1} B_2^T S_3^{-1} C_2) S_2^{-1} C_1 = 0, \\ A_{22} &= S_2^{-1} B_2^T S_3^{-1} C_2 S_2^{-1} C_1 A_1^{-1} B_1^T (-I + S_2^{-1} B_2^T S_3^{-1} C_2) \\ &\quad - 2S_2^{-1} B_2^T S_3^{-1} C_2 (I - S_2^{-1} B_2^T S_3^{-1} C_2) = 0, \\ A_{31} &= -S_3^{-1} C_2 S_2^{-1} C_1 (-I + A_1^{-1} B_1^T S_2^{-1} C_1) \\ &\quad + 2S_3^{-1} C_2 (-I + S_2^{-1} B_2^T S_3^{-1} C_2) S_2^{-1} C_1 = 0, \\ A_{32} &= -S_3^{-1} C_2 S_2^{-1} C_1 A_1^{-1} B_1^T (-I + S_2^{-1} B_2^T S_3^{-1} C_2) \\ &\quad + 2S_3^{-1} C_2 (I - S_2^{-1} B_2^T S_3^{-1} C_2) = 0, \\ A_{13} &= 0, \quad A_{23} = 0, \quad A_{33} = 0.\end{aligned}$$

We therefore conclude that

$$(\mathcal{T}_1 - I)(\mathcal{T}_1^2 - \mathcal{T}_1 - I)(\mathcal{T}_1^3 - \mathcal{T}_1^2 - 2\mathcal{T}_1 + I) = 0. \quad (2.7)$$

The proof is complete.  $\square$

We note that (2.7) can be factorized into distinct linear factors. Correspondingly, the roots of the above polynomial are

$$-1.2470, -0.6180, 0.4450, 1, 1.6180, 1.8019.$$

It follows that the eigenvalues of  $\mathcal{T}_1$  must be among these values. One can use the method of contradiction to prove that there is no polynomial of degree less than 6 such that  $\mathcal{T}_1$  satisfies the polynomial equation. Theorem 2.3 in [38] shows that if  $\mathcal{A}$  is a symmetric matrix, then the eigenvalues of the preconditioned system  $\mathcal{T}_1$  are

$$\sigma_p(\mathcal{T}_1) = \left\{ 2 \cos \left( \frac{2i-1}{2j+1} \pi \right) : j=1, \dots, 3, i=1, \dots, j \right\}.$$

Correspondingly, the condition number is

$$\kappa(\mathcal{T}_1) \leq \frac{\cos(\pi/7)}{\sin(\pi/14)} \approx 4.05.$$

Our conclusion for eigenvalues in Theorem 2.2 is consistent with that in [38]. However, the assumption that  $\mathcal{A}$  is symmetric and is removed in our analysis.

We also consider the following Schur complement based block-diagonal preconditioners:

$$\begin{aligned} \mathcal{P}_{D_2} &= \begin{bmatrix} A_1 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & -S_3 \end{bmatrix}, \quad \mathcal{P}_{D_3} = \begin{bmatrix} A_1 & 0 & 0 \\ 0 & -S_2 & 0 \\ 0 & 0 & S_3 \end{bmatrix}, \\ \mathcal{P}_{D_4} &= \begin{bmatrix} A_1 & 0 & 0 \\ 0 & -S_2 & 0 \\ 0 & 0 & -S_3 \end{bmatrix}. \end{aligned} \quad (2.8)$$

Furthermore, if  $A_2 = 0$  and  $A_3 = 0$ , we have the following conclusions: If the preconditioner is  $\mathcal{P}_{D_2}$ , then

$$\mathcal{T}_2 = \mathcal{P}_{D_2}^{-1} \mathcal{A} = \begin{bmatrix} I & A_1^{-1} B_1^T & 0 \\ S_2^{-1} C_1 & 0 & S_2^{-1} B_2^T \\ 0 & -S_3^{-1} C_2 & 0 \end{bmatrix},$$

satisfies the polynomial equation  $(\lambda - 1)(\lambda^2 - \lambda - 1)(\lambda^3 - \lambda^2 - 1) = 0$ . Its roots are

$$-0.618, -0.2328 + 0.7926i, -0.2328 - 0.7926i, 1, 1.4656, 1.618;$$

If the preconditioner is  $\mathcal{P}_{D_3}$ , then

$$\mathcal{T}_3 = \mathcal{P}_{D_3}^{-1} \mathcal{A} = \begin{bmatrix} I & A_1^{-1}B_1^T & 0 \\ -S_2^{-1}C_1 & 0 & -S_2^{-1}B_2^T \\ 0 & S_3^{-1}C_2 & 0 \end{bmatrix},$$

which satisfies the polynomial equation  $(\lambda - 1)(\lambda^2 - \lambda + 1)(\lambda^3 - \lambda^2 + 2\lambda - 1) = 0$ . Correspondingly, the roots are

$$0.5698, 1, 0.5 + 0.8660i, 0.5 - 0.8660i, 0.2151 + 1.3071i, 0.2151 - 1.3071i.$$

If the preconditioner is  $\mathcal{P}_{D_4}$ , then

$$\mathcal{T}_4 = \mathcal{P}_{D_4}^{-1} \mathcal{A} = \begin{bmatrix} I & A_1^{-1}B_1^T & 0 \\ -S_2^{-1}C_1 & 0 & -S_2^{-1}B_2^T \\ 0 & -S_3^{-1}C_2 & 0 \end{bmatrix},$$

which satisfies the polynomial equation  $(\lambda - 1)(\lambda^2 - \lambda + 1)(\lambda^3 - \lambda^2 + 1) = 0$ . The roots are

$$-0.7549, 1, 0.5 + 0.8660i, 0.5 - 0.8660i, 0.8774 + 0.7449i, 0.8774 - 0.7449i.$$

We highlight here that all eigenvalues of  $\mathcal{T}_3$  have positive real parts. Therefore,  $\mathcal{P}_{D_3}$  is a positively stable preconditioner for  $\mathcal{A}$ . We also comment here that when the  $A_1$  and the Schur complements  $S_2$  and  $S_3$  are replaced by their inexact versions, some analysis on the eigenvalue distribution of the preconditioned systems are obtained in [7].

### 3 Nested Schur complement based preconditioners for $n$ -tuple block tridiagonal problems

Now, we will discuss preconditioners for the  $n$ -tuple block tridiagonal problem of the following form:

$$\mathcal{A}_n = \begin{bmatrix} A_1 & B_1^T & & & \\ C_1 & -A_2 & \ddots & & \\ & \ddots & \ddots & B_{n-1}^T & \\ & & & C_{n-1} & (-1)^{n-1}A_n \end{bmatrix}. \quad (3.1)$$

Similar to the 3-by-3 block case, we define

$$S_{i+1} = A_{i+1} + C_i S_i^{-1} B_i^T, \quad i=1,2,\dots,n-1$$

with the initial setting  $S_1 = A_1$ . Here and hereafter, we assume that  $S_i, i=1,2,\dots,n$  are invertible. The block tridiagonal system can be written recursively as

$$\mathcal{A}_{i+1} = \begin{bmatrix} \mathcal{A}_i & \mathbb{B}_i^T \\ \mathbb{C}_i & (-1)^i A_{i+1} \end{bmatrix}.$$

Here,

$$\mathbb{B}_i^T = [0, \dots, 0, B_i]^T, \quad \mathbb{C}_i = [0, \dots, 0, C_i].$$

Then, following the notations introduced in [38],

$$\begin{aligned} \text{Schur}(\mathcal{A}_{i+1}) &= (-1)^i A_{i+1} - \mathbb{C}_i \mathcal{A}_i^{-1} \mathbb{B}_i^T \\ &= (-1)^i (A_{i+1} + C_i S_i^{-1} B_i^T), \end{aligned}$$

and

$$S_{i+1} = (-1)^i \text{Schur}(\mathcal{A}_{i+1}) = A_{i+1} + C_i S_i^{-1} B_i^T$$

with  $S_1 = A_1$ .

Similar to the factorization (2.1),  $\mathcal{A}_n$  can be factorized into

$$\mathcal{A}_n = \mathcal{L}_n \mathcal{D}_n \mathcal{U}_n, \quad (3.2)$$

where

$$\mathcal{D}_n = \begin{bmatrix} A_1 & 0 & & & \\ 0 & -S_2 & \ddots & & \\ & \ddots & \ddots & 0 & \\ & & & 0 & (-1)^{n-1} S_n \end{bmatrix}, \quad (3.3)$$

and

$$\begin{aligned} \mathcal{L}_n &= \begin{bmatrix} I & 0 & & & \\ C_1 A_1^{-1} & I & \ddots & & \\ & \ddots & \ddots & 0 & \\ & & (-1)^{n-2} C_{n-1} S_{n-1}^{-1} & I \end{bmatrix}, \\ \mathcal{U}_n &= \begin{bmatrix} I & A_1^{-1} B_1^T & & & \\ 0 & I & \ddots & & \\ & \ddots & \ddots & (-1)^{n-2} S_{n-1}^{-1} B_{n-1}^T & \\ & & 0 & I & \end{bmatrix}. \end{aligned}$$

**Theorem 3.1.** Assume that all  $S_i, i=1,2,3,\dots,n$  are invertible. Let the preconditioner be

$$\mathcal{P}_n = \begin{bmatrix} A_1 & & & & \\ C_1 & -S_2 & & & \\ & \ddots & \ddots & & \\ & & & C_{n-1} & (-1)^{n-1}S_n \end{bmatrix} = \mathcal{L}_n \mathcal{D}_n,$$

then  $\mathcal{P}_n^{-1} \mathcal{A}_n = \mathcal{U}_n$ , and  $\mathcal{P}_n^{-1} \mathcal{A}_n$  satisfies the polynomial  $(\lambda - 1)^n = 0$ . Furthermore, if  $\mathcal{A}_n$  is symmetric and  $A_i, i=1,2,\dots,n$  are symmetric positive definite, let the preconditioner be  $\mathcal{D}_n$ , then all eigenvalues of  $\mathcal{D}_n^{-1} \mathcal{A}_n$  are positive real.

*Proof.* Combining  $\mathcal{L}_n \mathcal{D}_n$  in (3.2), the first part of the conclusion is obvious. For the second part, noting that if  $\mathcal{A}_n$  is symmetric and  $A_i, i=1,2,\dots,n$ , are symmetric positive definite, then the matrix

$$\mathcal{S}_n = \begin{bmatrix} S_1 & & & & \\ & S_2 & & & \\ & & \ddots & & \\ & & & & S_n \end{bmatrix} \quad (3.4)$$

is symmetric positive definite [38]. We therefore can define  $\mathcal{S}_n^{1/2}$ , in which the  $i$ -th diagonal block is  $S_i^{1/2}$ . To prove the eigenvalues of  $\mathcal{D}_n^{-1} \mathcal{A}_n$  are positive real, we apply a similarity transform to  $\mathcal{D}_n^{-1} \mathcal{A}_n$ , that is, left multiplying by  $\mathcal{S}_n^{1/2}$  and right multiplying by  $\mathcal{S}_n^{-1/2}$  to  $\mathcal{D}_n^{-1} \mathcal{A}_n$ , we have

$$\begin{aligned} & \mathcal{S}_n^{1/2} \mathcal{D}_n^{-1} \mathcal{A}_n \mathcal{S}_n^{-1/2} \\ &= \begin{bmatrix} I & S_1^{-\frac{1}{2}} B_1^T S_2^{-\frac{1}{2}} & & & \\ -S_2^{-\frac{1}{2}} B_1 S_1^{-\frac{1}{2}} & S_2^{-\frac{1}{2}} A_2 S_2^{-\frac{1}{2}} & \ddots & & \\ & \ddots & \ddots & & (-1)^{n-2} S_{n-1}^{-\frac{1}{2}} B_{n-1}^T S_n^{-\frac{1}{2}} \\ & & & (-1)^{n-1} S_n^{-\frac{1}{2}} B_{n-1} S_{n-1}^{-\frac{1}{2}} & S_n^{-\frac{1}{2}} A_n S_n^{-\frac{1}{2}} \end{bmatrix}. \end{aligned}$$

We note that the Hermitian part of  $\mathcal{S}_n^{1/2} \mathcal{D}_n^{-1} \mathcal{A}_n \mathcal{S}_n^{-1/2}$  is

$$\frac{1}{2} \left[ \left( \mathcal{S}_n^{1/2} \mathcal{D}_n^{-1} \mathcal{A}_n \mathcal{S}_n^{-1/2} \right) + \left( \mathcal{S}_n^{1/2} \mathcal{D}_n^{-1} \mathcal{A}_n \mathcal{S}_n^{-1/2} \right)^T \right] = \begin{bmatrix} I & 0 & & & \\ 0 & S_2^{-\frac{1}{2}} A_2 S_2^{-\frac{1}{2}} & \ddots & & \\ & \ddots & \ddots & & 0 \\ & & & & \\ 0 & S_n^{-\frac{1}{2}} A_n S_n^{-\frac{1}{2}} & & & \end{bmatrix}$$

is symmetric positive definite. Therefore, we conclude that all eigenvalues of  $\mathcal{D}_n^{-1}\mathcal{A}_n$  are positive real.  $\square$

**Remark 3.1.** In [38],  $\mathcal{S}_n$  is taken as a preconditioner of the block tridiagonal system  $\mathcal{A}_n$ . The advantage is that  $\mathcal{S}_n$  is SPD. However, the eigenvalues of  $\mathcal{S}_n^{-1}\mathcal{A}_n$  may not be positive real. In comparison,  $\mathcal{D}_n$  is a positively stable preconditioner for  $\mathcal{A}_n$ .

**Theorem 3.2.** *Assume that  $A_i = 0$  for  $i = 2, 3, \dots, n$  in  $\mathcal{A}_n$ . If the Schur complement based preconditioner is  $\mathcal{D}_n$  in (3.3), then the eigenvalues of  $\mathcal{T}_n = \mathcal{D}_n^{-1}\mathcal{A}_n$  are the roots of polynomial*

$$\bar{p}_i = \lambda \bar{p}_{i-1} + \bar{p}_{i-2}, \quad i = 2, 3, 4, \dots, n \quad (3.5)$$

with  $\bar{p}_0(\lambda) = 1$ ,  $\bar{p}_1(\lambda) = \lambda - 1$ . Moreover, all the roots of  $\bar{p}_i$ ,  $i = 1, 2, \dots, n$ , lie in the right half-plane. If  $\mathcal{T}_n = \mathcal{D}_n^{-1}\mathcal{A}_n$  is further assumed to be diagonalizable, then  $\mathcal{T}_n$  satisfies the polynomial equation  $\prod_{i=0}^n \bar{p}_i = 0$ .

*Proof.* To verify the first part of the theorem, we let

$$\mathcal{C}_i = S_{i+1}^{-1}C_i, \quad \mathcal{B}_i^T = S_i^{-1}B_i^T, \quad i = 1, 2, \dots, n-1.$$

Because  $A_i = 0$ ,  $i = 2, \dots, n$ , we see that  $\mathcal{C}_i \mathcal{B}_i^T = I$  and the preconditioned system  $\mathcal{T}_n$  degenerates to

$$\mathcal{T}_n = \mathcal{D}_n^{-1}\mathcal{A}_n = \begin{bmatrix} I & \mathcal{B}_1^T \\ -\mathcal{C}_1 & 0 & \ddots \\ \ddots & \ddots & (-1)^{n-2}\mathcal{B}_{n-1}^T \\ (-1)^{n-1}\mathcal{C}_{n-1} & 0 \end{bmatrix}. \quad (3.6)$$

Let  $\lambda$  be an eigenvalue and  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  be the corresponding eigenvector, we have  $\mathcal{T}_n \mathbf{x} = \lambda \mathbf{x}$ . More clearly,

$$\begin{aligned} x_1 + \mathcal{B}_1^T x_2 &= \lambda x_1, \\ \mathcal{C}_1 x_1 + \mathcal{B}_2^T x_3 &= -\lambda x_2, \\ &\vdots \\ \mathcal{C}_{n-2} x_{n-2} + \mathcal{B}_{n-1}^T x_n &= (-1)^{n-2} \lambda x_{n-1}, \\ \mathcal{C}_{n-1} x_{n-1} &= (-1)^{n-1} \lambda x_n. \end{aligned} \quad (3.7)$$

From the first equation of (3.7),

$$\mathcal{B}_1^T x_2 = \bar{p}_1(\lambda) x_1, \quad \text{where} \quad \bar{p}_1(\lambda) = \lambda - 1.$$

Setting  $x_2 = x_3 = \dots = x_n = 0$ , and  $x_1 \neq 0$ , we note that  $\lambda_{11} = 1$ , the root of  $\bar{p}_1(\lambda)$ , is an eigenvalue. If  $\lambda \neq \lambda_{11}$ , from the second equation and replacing  $x_1$  by  $\mathcal{B}_1^T x_2 / \bar{p}_1(\lambda)$ , we then obtain

$$\mathcal{B}_2^T x_3 = -\mathcal{C}_1 x_1 - \lambda x_2 = -\frac{1}{\bar{p}_1(\lambda)} \mathcal{C}_1 \mathcal{B}_1^T x_2 - \lambda x_2 = -\frac{1}{\bar{p}_1(\lambda)} x_2 - \lambda x_2 = -R_2(\lambda) x_2,$$

where

$$R_2(\lambda) = \lambda + \frac{1}{\bar{p}_1(\lambda)} = \frac{\bar{p}_2(\lambda)}{\bar{p}_1(\lambda)}$$

with

$$\bar{p}_2(\lambda) = \lambda \bar{p}_1(\lambda) + 1 = \lambda \bar{p}_1(\lambda) + \bar{p}_0(\lambda).$$

By setting  $x_3 = \dots = x_n = 0$ , we see that the two roots of the polynomial  $\bar{p}_2(\lambda)$ , denoted as  $\lambda_{21}$  and  $\lambda_{22}$ , are eigenvalues. Repeating this procedure leads to

$$\mathcal{B}_j^T x_{j+1} = (-1)^{j-1} R_j(\lambda) x_j, \quad j = 2, \dots, n-1,$$

and

$$0 = (-1)^{n-1} R_n(\lambda) x_n \quad \text{with} \quad R_j(\lambda) = \frac{\bar{p}_j(\lambda)}{\bar{p}_{j-1}(\lambda)},$$

where the polynomials  $\bar{p}_j(\lambda)$  are recursively given by

$$\bar{p}_0(\lambda) = 1, \quad \bar{p}_1(\lambda) = \lambda - 1, \quad \bar{p}_{i+1}(\lambda) = \lambda \bar{p}_i(\lambda) + \bar{p}_{i-1}(\lambda), \quad i \geq 1.$$

Thus, the eigenvalues of  $\mathcal{T}_n$  are the roots of the polynomials  $\bar{p}_1(\lambda), \bar{p}_2(\lambda), \dots, \bar{p}_n(\lambda)$ .

For proving the second part of the conclusion, we note that all eigenvalues of  $\mathcal{D}_n^{-1} \mathcal{A}_n$  are the roots of  $\bar{p}_i$ ,  $i = 1, 2, \dots, n$ . We are going to show that all roots of the polynomials  $\bar{p}_i$ ,  $i = 1, 2, \dots, n$ , lie on the right half-plane by using the Routh-Hurwitz stability criterion [2, 20].

The stability of a polynomial can be determined by the Routh-Hurwitz stability criterion which suggests using a tabular method [2]: For a general  $k$ -th degree polynomial

$$p_k(\lambda) = a_k \lambda^k + a_{k-1} \lambda^{k-1} + \dots + a_1 \lambda + a_0$$

its Routh table is a matrix of  $k+1$  rows, which has the following structure:

$$R_{T_k} = \begin{bmatrix} r_{01} & r_{02} & r_{03} & \cdots \\ r_{11} & r_{12} & r_{13} & \cdots \\ r_{21} & r_{22} & r_{23} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (3.8)$$

In (3.8), the first two rows, which come from the coefficients directly, are ordered as

$$(r_{0j}) = (a_k, a_{k-2}, a_{k-4}, \dots), \quad (r_{1j}) = (a_{k-1}, a_{k-3}, a_{k-5}, \dots). \quad (3.9)$$

Other entries are computed as follows:

$$r_{ij} = \frac{-1}{r_{i-1,1}} \det \begin{bmatrix} r_{i-2,1} & r_{i-2,j+1} \\ r_{i-1,1} & r_{i-1,j+1} \end{bmatrix}, \quad i=2,3,4,\dots$$

When the table is completed, the number of sign changes in the first column is the number of roots in the right half-plane [2]. Note that the entries of the Routh table are uniquely determined once the first two rows are fixed, i.e., the Routh table is uniquely dependent on the coefficients of  $p_k$ .

For the polynomial  $\bar{p}_k$  generated by (3.5), its key coefficients (for the leading terms and the constant) are

$$\bar{a}_k = 1, \quad \bar{a}_{k-1} = -1, \quad \bar{a}_0 = (-1)^k. \quad (3.10)$$

Such a conclusion can be proved by the method of mathematical induction by using the recurrence formula (3.5). From (3.10), in general, we have

$$\bar{p}_k = \lambda^k - \lambda^{k-1} + \bar{a}_{k-2}\lambda^{k-2} + \bar{a}_{k-3}\lambda^{k-3} + \bar{a}_{k-4}\lambda^{k-4} + \cdots + \bar{a}_2\lambda^2 + \bar{a}_1\lambda + (-1)^k.$$

For completing the Routh table of  $\bar{p}_k$ , let us write out the expression of  $\bar{p}_{k-1}$  in terms of the coefficients of  $\bar{p}_k$ . From the recurrence formula (3.5), we note that the expression of  $\bar{p}_{k-1}$  will have small differences for the case  $k$  odd and the case  $k$  even

$$\begin{aligned} \bar{p}_{k-1} = & \lambda^{k-1} - \lambda^{k-2} - \bar{a}_{k-3}\lambda^{k-3} - (\bar{a}_{k-4} + \bar{a}_{k-5})\lambda^{k-4} - \bar{a}_{k-5}\lambda^{k-5} \\ & - (\bar{a}_{k-6} + \bar{a}_{k-7})\lambda^{k-6} \cdots - (\bar{a}_5 + \bar{a}_4)\lambda^5 - \bar{a}_4\lambda^4 - (\bar{a}_3 + \bar{a}_2)\lambda^3 \\ & - \bar{a}_2\lambda^2 - (\bar{a}_1 + \bar{a}_0)\lambda - \bar{a}_0, \quad \text{if } k \text{ is odd,} \end{aligned} \quad (3.11)$$

$$\begin{aligned} \bar{p}_{k-1} = & \lambda^{k-1} - \lambda^{k-2} - \bar{a}_{k-3}\lambda^{k-3} - (\bar{a}_{k-4} + \bar{a}_{k-5})\lambda^{k-4} - \bar{a}_{k-5}\lambda^{k-5} \\ & - (\bar{a}_{k-6} + \bar{a}_{k-7})\lambda^{k-6} \cdots - \bar{a}_5\lambda^5 - (\bar{a}_4 + \bar{a}_3)\lambda^4 - \bar{a}_3\lambda^3 \\ & - (\bar{a}_2 + \bar{a}_1)\lambda^2 - \bar{a}_1\lambda - 1, \quad \text{if } k \text{ is even.} \end{aligned} \quad (3.12)$$

For ease of presentation, let us assume that  $k$  is even (then  $k-1$  and  $k+1$  are odd. The case  $k$  is odd can be proved similarly by slightly changing the notations in the following arguments). We denote the matrix  $R_{T_k}$  as the Routh table of polynomial  $\bar{p}_k$ . The first two rows of  $R_{T_k}$  are

$$\begin{bmatrix} \bar{a}_k = 1 & \bar{a}_{k-2} & \bar{a}_{k-4} & \cdots & \bar{a}_2 & \bar{a}_0 = 1 \\ \bar{a}_{k-1} = -1 & \bar{a}_{k-3} & \bar{a}_{k-5} & \cdots & \bar{a}_1 & 0 \end{bmatrix}. \quad (3.13)$$

Then, the first two rows of  $R_{T_{k-1}}$  are

$$\begin{bmatrix} 1 & -\bar{a}_{k-3} & -\bar{a}_{k-5} & \cdots & -\bar{a}_1 \\ -1 & -(\bar{a}_{k-4} + \bar{a}_{k-5}) & -(\bar{a}_{k-6} + \bar{a}_{k-7}) & \cdots & -1 \end{bmatrix}, \quad (3.14)$$

the first two rows of  $R_{T_{k+1}}$  are

$$\begin{bmatrix} 1 & \bar{a}_{k-2} - \bar{a}_{k-1} & \bar{a}_{k-4} - \bar{a}_{k-3} & \cdots & \bar{a}_2 - \bar{a}_3 & \bar{a}_0 - \bar{a}_1 \\ -\bar{a}_k = -1 & -\bar{a}_{k-2} & -\bar{a}_{k-4} & \cdots & -\bar{a}_2 & -\bar{a}_0 \end{bmatrix}, \quad (3.15)$$

and  $R_{T_{k+1}}$  satisfy the following relationship:

$$R_{T_{k+1}} = \begin{bmatrix} 1 & \bar{a}_{k-2} - \bar{a}_{k-1} & \bar{a}_{k-4} - \bar{a}_{k-3} & \cdots & \bar{a}_2 - \bar{a}_3 & \bar{a}_0 - \bar{a}_1 \\ & & -R_{T_k} & & & \end{bmatrix}. \quad (3.16)$$

One only needs to verify that the derivation from (3.13) to (3.15) is correct. To check this, by using (3.5), there holds

$$\begin{aligned} \bar{p}_{k+1} &= \lambda \bar{p}_k + \bar{p}_{k-1} \\ &= \lambda(\lambda^k - \lambda^{k-1} + \bar{a}_{k-2}\lambda^{k-2} + \bar{a}_{k-3}\lambda^{k-3} + \bar{a}_{k-4}\lambda^{k-4} \dots + \bar{a}_2\lambda^2 + \bar{a}_1\lambda + \bar{a}_0) + \lambda^{k-1} \\ &\quad - \lambda^{k-2} - \bar{a}_{k-3}\lambda^{k-3} - (\bar{a}_{k-4} + \bar{a}_{k-5})\lambda^{k-4} - \bar{a}_{k-5}\lambda^{k-5} \dots - (\bar{a}_{k-6} + \bar{a}_{k-7})\lambda^{k-6} \\ &\quad - (\bar{a}_5 + \bar{a}_4)\lambda^5 - \bar{a}_4\lambda^4 - (\bar{a}_3 + \bar{a}_2)\lambda^3 - \bar{a}_2\lambda^2 - (\bar{a}_1 + \bar{a}_0)\lambda - \bar{a}_0 \\ &= \lambda^{k+1} - \lambda^k + (\bar{a}_{k-2} + 1)\lambda^{k-1} + (\bar{a}_{k-3} - 1)\lambda^{k-2} + (\bar{a}_{k-4} - \bar{a}_{k-3})\lambda^{k-3} - \bar{a}_{k-4}\lambda^{k-4} \\ &\quad + (\bar{a}_{k-6} - \bar{a}_{k-5})\lambda^{k-5} - \bar{a}_{k-6}\lambda^{k-6} \dots + (\bar{a}_4 - \bar{a}_3)\lambda^2 - \bar{a}_3\lambda^3 + (\bar{a}_1 - \bar{a}_2)\lambda^2 - \bar{a}_1\lambda + 1. \end{aligned}$$

As  $k$  is even,  $k+1$  is odd, the expression of  $\bar{p}_{k+1}$  satisfies the first equation in (3.11), and the first two rows of the Routh table of  $\bar{p}_{k+1}$  are

$$\begin{bmatrix} 1 & \bar{a}_{k-2} - \bar{a}_{k-1} & \bar{a}_{k-4} - \bar{a}_{k-3} & \cdots & \bar{a}_2 - \bar{a}_3 & \bar{a}_0 - \bar{a}_1 \\ -\bar{a}_k = -1 & -\bar{a}_{k-2} & -\bar{a}_{k-4} & \cdots & -\bar{a}_2 & -\bar{a}_0 \end{bmatrix}.$$

Thus, (3.15) is verified. As the entries of the Routh table are uniquely determined by the first two rows, the rest of the verifications are straightforward: for verifying (3.16), one can directly check the entries row-by-row by using (3.9). Actually, the entries of the 3-rd row of  $R_{T_{k+1}}$  are

$$\begin{aligned} r_{21} &= \frac{-1}{-1} \det \begin{bmatrix} 1 & \bar{a}_{k-2} - \bar{a}_{k-1} \\ -1 & -\bar{a}_{k-2} \end{bmatrix} = -\bar{a}_{k-1} = 1, \\ r_{22} &= \frac{-1}{-1} \det \begin{bmatrix} 1 & \bar{a}_{k-4} - \bar{a}_{k-3} \\ -1 & -\bar{a}_{k-4} \end{bmatrix} = -\bar{a}_{k-3}, \dots, \\ r_{2j} &= \frac{-1}{-1} \det \begin{bmatrix} 1 & \bar{a}_{k-2j} - \bar{a}_{k-2j+1} \\ -1 & -\bar{a}_{k-2j} \end{bmatrix} = -\bar{a}_{k-2j+1} \quad \text{for a general } j. \end{aligned}$$

By comparing these entries with the 2nd row of (3.13), one sees that (3.16) is verified.

We conclude that the entries of the first row of  $R_{T_k}$  change values from 1 to  $-1$  alternatingly. Therefore, the number of sign changes in the first column is  $k$ . Thus, according to the Routh-Hurwitz stability criterion [2, 20] and the fact that there is no zero root, the number of roots of  $\bar{p}_k$  which are lying in the right half-plane equals  $k$ . Therefore, all the roots of  $\bar{p}_i$ ,  $i=1,2,\dots,n$ , lie in the right half-plane.

If  $\mathcal{T}_n = \mathcal{D}_n^{-1} \mathcal{A}_n$  is further assumed to be diagonalizable, then there exists an invertible matrix, say  $\mathcal{Q}_n$ , such that  $\mathcal{Q}_n^{-1} \mathcal{T}_n \mathcal{Q}_n$  is a diagonal matrix with the diagonal entries being the eigenvalues. By a direct verification,  $\mathcal{T}_n$  satisfies the polynomial equation  $\prod_{i=0}^n \bar{p}_i = 0$ .  $\square$

**Remark 3.2.** We highlight here that in the assumption of Theorem 3.2, we do not require that  $\mathcal{A}_n$  is symmetric. Moreover, under the same assumption of Theorem 3.2, if  $\mathcal{S}_n$  in (3.4) is the preconditioner, then the eigenvalues of  $\mathcal{S}_n^{-1} \mathcal{A}_n$  are the roots of the polynomial of the form  $\prod_{i=0}^n \tilde{p}_i$ , where  $\tilde{p}_0(\lambda) = 1$ ,  $\tilde{p}_1(\lambda) = \lambda - 1$  and

$$\tilde{p}_i = \lambda \tilde{p}_{i-1} - \tilde{p}_{i-2}, \quad i = 2, 3, 4, \dots, n.$$

The proof of this conclusion is the same as that provided in [38], which is for the symmetric case. However,  $\mathcal{S}_n$  is not a positively stable preconditioner. In comparison, we show that  $\mathcal{D}_n$  is a positively stable preconditioner for  $\mathcal{A}_n$  no matter whether it is symmetric or not.

**Remark 3.3.** We comment here that if  $\mathcal{A}_n$  is symmetric, the proof for verifying  $\mathcal{D}_n$  is positively stable can be carried out similarly to that for Theorem 3.1. More

clearly, we left multiply by  $\mathcal{S}_n^{1/2}$  and right multiply by  $\mathcal{S}_n^{-1/2}$  to  $\mathcal{D}_n^{-1}\mathcal{A}_n$ , and denote  $\mathcal{H}_n = \mathcal{S}_n^{1/2}\mathcal{D}_n^{-1}\mathcal{A}_n\mathcal{S}_n^{-1/2}$ . Then, we have

$$\mathcal{H}_n = \begin{bmatrix} I & \mathcal{S}_1^{-\frac{1}{2}}B_1^T\mathcal{S}_2^{-\frac{1}{2}} & & & \\ -\mathcal{S}_2^{-\frac{1}{2}}B_1\mathcal{S}_1^{-\frac{1}{2}} & 0 & \ddots & & \\ & \ddots & \ddots & (-1)^{n-2}\mathcal{S}_{n-1}^{-\frac{1}{2}}B_{n-1}^T\mathcal{S}_n^{-\frac{1}{2}} & \\ & & (-1)^{n-1}\mathcal{S}_n^{-\frac{1}{2}}B_{n-1}\mathcal{S}_{n-1}^{-\frac{1}{2}} & 0 & \end{bmatrix}.$$

We note that the Hermitian part of  $\mathcal{S}_n^{1/2}\mathcal{D}_n^{-1}\mathcal{A}_n\mathcal{S}_n^{-1/2}$  is

$$\frac{1}{2} \left[ (\mathcal{S}_n^{\frac{1}{2}}\mathcal{D}_n^{-1}\mathcal{A}_n\mathcal{S}_n^{-\frac{1}{2}}) + (\mathcal{S}_n^{\frac{1}{2}}\mathcal{D}_n^{-1}\mathcal{A}_n\mathcal{S}_n^{-\frac{1}{2}})^T \right] = \begin{bmatrix} I & 0 & & & \\ 0 & 0 & \ddots & & \\ & \ddots & \ddots & 0 & \\ & & 0 & 0 & \end{bmatrix}.$$

Let us denote  $\lambda$  as an eigenvalue of  $\mathcal{H}_n$ ,  $\bar{\lambda}$  as its conjugate, and  $\mathbf{x}$  as the corresponding eigenvector. As  $\mathcal{H}_n$  is invertible,  $\lambda$  is nonzero and  $\mathbf{x}$  is not a zero vector. We have

$$\mathbf{x}^T \mathcal{H}_n \mathbf{x} = \lambda \|\mathbf{x}\|^2, \quad \mathbf{x}^T \mathcal{H}_n^T \mathbf{x} = \bar{\lambda} \|\mathbf{x}\|^2.$$

Summing the two equations together and noting from the expression of  $\mathcal{H}_n$ , we have

$$\begin{aligned} \mathbf{x}^T (\mathcal{H}_n + \mathcal{H}_n^T) \mathbf{x} &= 2\operatorname{Re}(\lambda) \|\mathbf{x}\|^2, \\ 2\|\mathbf{x}\|^2 &= 2\operatorname{Re}(\lambda) \|\mathbf{x}\|^2. \end{aligned} \tag{3.17}$$

Here,  $x_1$  is the first component of  $\mathbf{x}$  and  $\operatorname{Re}(\lambda)$  is the real part of  $\lambda$ . We assert that  $x_1$  is nonzero. This conclusion can be proved by the method of contradiction. If  $x_1$  is 0, starting with the first equation of (3.7), we have  $x_2 \in \operatorname{Ker}(\mathcal{B}_1^T)$ . However,  $\mathcal{B}_1^T$  must be of full column rank because  $\mathcal{B}_1\mathcal{B}_1^T = I$ . It follows that  $x_2$  is 0. Repeating this procedure as presented in (3.7), and noting that  $\mathcal{B}_i^T$  is of full column rank as  $\mathcal{B}_i\mathcal{B}_i^T = I$ , we will have  $x_3 = \dots = x_n = 0$ . It contradicts the fact that  $\mathbf{x}$  is nonzero. Therefore, from (3.17), we conclude that  $\operatorname{Re}(\lambda) > 0$ .

## 4 Additive Schur complement based preconditioners for twofold saddle point problems

In this section, we discuss how to design preconditioners when the twofold saddle point system is permuted into (1.2). Such a form of the linear system arises naturally from domain decomposition methods with two subdomains [39]. Without confusion, we introduce simplified notation

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_3 \end{bmatrix}, \quad B^T = \begin{bmatrix} B_1^T \\ C_2 \end{bmatrix}, \quad C = [C_1 \ B_2^T]. \quad (4.1)$$

Here,  $A_3$  can degenerate to 0. If  $A_1$  and  $A_3$  are invertible, then  $A$  is invertible, and the Schur complement simply reads as

$$S = A_2 + CA^{-1}B^T.$$

Furthermore, if  $S$  is assumed to be invertible, then the following results hold [19, 25].

**Proposition 4.1.** *If  $A$  and  $S$  are invertible and the preconditioner is*

$$\mathcal{Q}_1 = \begin{bmatrix} A & 0 \\ C & -S \end{bmatrix},$$

*then*

$$\mathcal{Q}_1^{-1} \mathcal{A} = \begin{bmatrix} I & A^{-1}B^T \\ 0 & I \end{bmatrix},$$

*which satisfies the polynomial equation  $(\lambda - 1)^2 = 0$ . Moreover, if  $A_2 = 0$  and the preconditioner is*

$$\mathcal{Q}_{D_1} = \begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix},$$

*then*

$$\mathcal{T}_{Q_{D_1}} = \mathcal{Q}_{D_1}^{-1} \mathcal{A} = \begin{bmatrix} I & A^{-1}B^T \\ S^{-1}C & 0 \end{bmatrix},$$

*which satisfies the polynomial equation  $(\lambda(\lambda - 1))^2 = \lambda(\lambda - 1)$ . The non-zero eigenvalues of  $\mathcal{T}_{Q_{D_1}}$  are*

$$1, \frac{1 \pm \sqrt{5}}{2}.$$

The proof of this proposition can be found in [25, 31].  
Based on the formulation (4.1), if the preconditioner is

$$\mathcal{Q}_2 = \begin{bmatrix} A & 0 \\ C & S \end{bmatrix},$$

then

$$\mathcal{Q}_2^{-1} \mathcal{A} = \begin{bmatrix} I & A^{-1}B^T \\ 0 & -I \end{bmatrix},$$

which satisfies the polynomial equation  $(\lambda - 1)(\lambda + 1) = 0$ . Moreover, if  $A_2 = 0$  and the preconditioner is

$$\mathcal{Q}_{D_2} = \begin{bmatrix} A & 0 \\ 0 & -S \end{bmatrix},$$

then

$$\mathcal{T}_{Q_{D_2}} = \mathcal{Q}_{D_2}^{-1} \mathcal{A} = \begin{bmatrix} I & A^{-1}B^T \\ -S^{-1}C & 0 \end{bmatrix},$$

which satisfies the polynomial equation  $(\lambda(\lambda - 1))^2 = -\lambda(\lambda - 1)$ . The non-zero roots are

$$1, \frac{1 \pm \sqrt{3}i}{2}.$$

We note that the polynomial equations for  $\mathcal{Q}_i$  and  $\mathcal{Q}_{D_i}$  have a lower degree than  $\mathcal{P}_i$  and  $\mathcal{P}_{D_i}$  discussed in Section 2. From a purely algebraic point of view, if the GMRES method is applied,  $\mathcal{Q}_1$  is more favored than  $\mathcal{P}_1$ ,  $\mathcal{Q}_{D_2}$  is more favored than  $\mathcal{P}_{D_i}$ ,  $i = 1, 2, 3, 4$ . However, on the other hand, it might be not easy to directly derive a proper approximation for the additive Schur complement for some specific problems [30, 39]. For the nested Schur complement based preconditioners, they actually provide a recursive approach for solving block-tridiagonal systems.

## 5 Additive Schur complement based preconditioners from the viewpoint of domain decomposition

In this section, we generalize the discussion of Section 4 to the  $n$ -tuple case. We consider a system

$$\mathcal{A}_n = \begin{bmatrix} A_1 & 0 & \dots & B_1^T \\ 0 & -A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & B_{n-1}^T \\ C_1 & \dots & C_{n-1} & (-1)^{n-1} A_n \end{bmatrix}. \quad (5.1)$$

Formally, the system matrix (5.1) arises naturally from the domain decomposition method for an elliptic problem with  $n-1$  subdomains. Like in Section 4, one can define  $A, C, B^T$ , and  $D$ . The additive Schur complement preconditioner and the corresponding theory in Section 4 can still be applied. The resulting preconditioned system satisfies a polynomial of degree 2 (block triangular preconditioner) or 4 (block-diagonal preconditioner).

For some special forms of (5.1), for example, 3-by-3 block systems, it is possible to reorder [30, 37] the interior and interface unknowns subdomain by subdomain in a sequential manner and reduce them to the block tridiagonal systems discussed in Section 3. If an additive Schur complement based preconditioner is applied, then the degree of the polynomial equation that the preconditioned system satisfies is 2 (block triangular preconditioners) or 4 (block-diagonal preconditioners). See [25, 31] for the proof. In contrast, the preconditioners based on the nested Schur complement satisfy polynomials with degrees may be as high as  $n$  (block triangular) or  $n!$  (block-diagonal). Therefore, an additive Schur complement based preconditioner is more favored than a nested Schur complement based preconditioner. The Schur complement of (5.1), which corresponds to the Steklov-Poincaré operator [30, 39], is an additive type. Algebraically inverting the Schur complement directly is typically challenging, as there is no reduction in operational costs. Instead, various preconditioning techniques such as Dirichlet-Neumann, (balancing) Neumann-Neumann, fractional Sobolev norm-based preconditioners, vertex space preconditioners, and others can be developed, particularly for elliptic problems. For more detailed information and discussions, interested readers are directed to [30, 39]. For some specific PDE problems, for example, the Biot model [26, 33] or models from fluid dynamics [8, 10, 21], it is possible to derive an analytic approximation of the additive type Schur complement from a Fourier analysis point of view [8, 11]. In this work, for a special 3-field reformulation of the Biot model, based on its block tridiagonal form, we design nested Schur complement based preconditioners. In our experiments, the Schur complements are firstly approximated by using a Fourier analysis derivation, then are further approximated by using incomplete Cholesky factorizations. The detailed numerical experiments are provided in Section 6.

If  $n > 3$ , we realize that not all block tridiagonal systems can be directly permuted into the domain decomposition form (5.1), although the number of non-

zero blocks of (5.1) and that of (3.1) are the same for the non-degenerate cases. Therefore, it is also important to investigate the nested Schur complement based preconditioners.

## 6 Numerical experiments based on a 3-field formulation of the Biot model

A 3-field formulation of the Biot model reads as [26,33]

$$-2\mu \operatorname{div}(\boldsymbol{\epsilon}(\boldsymbol{u})) + \nabla \xi = \boldsymbol{f}, \quad (6.1)$$

$$-\operatorname{div} \boldsymbol{u} - \frac{1}{\lambda} \xi + \frac{\alpha}{\lambda} p = 0, \quad (6.2)$$

$$\left( \left( c_0 + \frac{\alpha^2}{\lambda} \right) p - \frac{\alpha}{\lambda} \xi \right)_t - K \operatorname{div}(\nabla p - \rho_f \boldsymbol{g}) = Q_s. \quad (6.3)$$

Here,  $\boldsymbol{u}$  is the displacement,  $p$  is the fluid pressure, and  $\xi = -\alpha p + \lambda \operatorname{div} \boldsymbol{u}$  is the total pressure with  $\alpha$  being Biot constant,  $K$  is the permeability,  $\boldsymbol{f}$  is the body force,  $Q_s$  is the source term

$$\boldsymbol{\epsilon}(\boldsymbol{u}) := \frac{1}{2} (\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T),$$

and

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \quad \mu = \frac{E}{2(1+\nu)}. \quad (6.4)$$

The boundary conditions are

$$\boldsymbol{u} = \mathbf{0} \quad \text{on } \Gamma_d, \quad (6.5)$$

$$\boldsymbol{\sigma}(\boldsymbol{u}) \boldsymbol{n} - \alpha p \boldsymbol{n} = \boldsymbol{h} \quad \text{on } \Gamma_t, \quad (6.6)$$

$$p = 0 \quad \text{on } \Gamma_p, \quad (6.7)$$

$$K(\nabla p - \rho_f \boldsymbol{g}) \cdot \boldsymbol{n} = g_2 \quad \text{on } \Gamma_f. \quad (6.8)$$

The measures of the Dirichlet boundary  $\Gamma_d$  and  $\Gamma_p$  are assumed to be greater than 0. For the wellposedness of the Biot problem with the above boundary conditions, we refer the readers to [26,33]. Without loss of generality, the Dirichlet conditions in (6.5)-(6.8) are assumed to be homogeneous. The initial conditions are

$$\boldsymbol{u}(0) = \boldsymbol{u}_0 \quad \text{and} \quad p(0) = p_0. \quad (6.9)$$

We apply the Taylor-Hood element pair for the discretization of  $\boldsymbol{u}$  and  $\xi$ , and  $P_1$  elements for  $p$ . The corresponding finite element spaces are denoted as  $V^h$ ,  $Q^h$ ,

and  $W^h$ . We apply a backward Euler scheme with a time stepsize  $\Delta t$  for (6.3). After the time and spatial discretizations, the resulting linear system is of the following form:

$$Ax = \begin{bmatrix} A_u & B_{u\xi}^T & 0 \\ B_{u\xi} & -A_\xi & B_{\xi p}^T \\ 0 & B_{\xi p} & A_p \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{\xi} \\ \bar{p} \end{bmatrix} = b. \quad (6.10)$$

Here,  $\bar{u}$ ,  $\bar{\xi}$ , and  $\bar{p}$  are unknowns associated with  $u$ ,  $\xi$ , and  $p$  respectively,  $A_u$ ,  $B_{u\xi}$ ,  $A_\xi$ ,  $B_{\xi p}$ , and  $A_p$  are the matrices resulted from the following bilinear forms:

$$\begin{aligned} a_u(u, v) &= 2\mu(\epsilon(u), \epsilon(v)), & b_{u\xi}(u, \phi) &= -(\phi, \operatorname{div} u), \\ a_\xi(\xi, \phi) &= \frac{1}{\lambda}(\xi, \phi), & b_{\xi p}(\xi, \psi) &= \frac{\alpha}{\lambda}(\xi, \psi), \\ a_p(p, \psi) &= -\left(\left(c_0 + \frac{\alpha^2}{\lambda}\right)p, \psi\right) - \Delta t K(\nabla p, \nabla \psi). \end{aligned}$$

We comment here that  $A_p$  is negative definite based on the weak form it represents. Nevertheless, the presentation in this section is consistent with the notations in the previous sections.

Applying a Fourier analysis derivation [8, 9, 11] firstly, we note that the differential operator for solving (6.1)-(6.9) is

$$\mathcal{A} = \begin{bmatrix} -2\mu\partial_x^2 - \mu\partial_y^2 & -\mu\partial_y\partial_x & \partial_x & 0 \\ -\mu\partial_y\partial_x & -\mu\partial_x^2 - 2\mu\partial_y^2 & \partial_y & 0 \\ -\partial_x & -\partial_y & -1/\lambda & \alpha/\lambda \\ 0 & 0 & \alpha/\lambda & -(c_0 + \alpha^2/\lambda) + K\Delta t\Delta \end{bmatrix}.$$

The Fourier mode of  $\hat{\mathcal{A}}$  is

$$\hat{\mathcal{A}} = \begin{bmatrix} 2\mu k^2 + \mu l^2 & \mu k l & i k & 0 \\ \mu k l & 2\mu l^2 + \mu k^2 & i l & 0 \\ -i k & -i l & -1/\lambda & \alpha/\lambda \\ 0 & 0 & \alpha/\lambda & -(c_0 + \alpha^2/\lambda) - K\Delta t(k^2 + l^2) \end{bmatrix}.$$

Thus, the Fourier mode for  $S_\xi = A_\xi + B_{u\xi} A_u^{-1} B_{u\xi}^T$  is

$$\hat{S}_\xi = \frac{1}{\lambda} + [-i k - i l] \begin{bmatrix} 2\mu k^2 + \mu l^2 & \mu k l \\ \mu k l & 2\mu l^2 + \mu k^2 \end{bmatrix}^{-1} \begin{bmatrix} i k \\ i l \end{bmatrix} = \frac{1}{\lambda} + \frac{1}{2\mu}.$$

Then, the Fourier mode for  $S_p = A_p + B_{\xi p} S_{\xi}^{-1} B_{\xi p}^T$  is

$$\begin{aligned}\hat{S}_p &= -\left(c_0 + \frac{\alpha^2}{\lambda}\right) - K\Delta t(k^2 + l^2) + \frac{\alpha}{\lambda} \left(\frac{1}{\lambda} + \frac{1}{2\mu}\right)^{-1} \frac{\alpha}{\lambda} \\ &= -\left(c_0 + \frac{\alpha^2}{\lambda}\right) - K\Delta t(k^2 + l^2) + \frac{2\mu\alpha^2}{\lambda(\lambda+2\mu)} \\ &= -\left(c_0 + \frac{\alpha^2}{2\mu+\lambda}\right) - K\Delta t(k^2 + l^2).\end{aligned}$$

From the above Fourier analysis, it is easy to see that  $\hat{S}_{\xi}$  and  $\hat{S}_p$  correspond to the following matrices:

$$S_{\xi} = \left(\frac{1}{\lambda} + \frac{1}{2\mu}\right) M_{\xi} \quad \text{and} \quad S_p = A_p + \frac{2\mu\alpha^2}{\lambda(\lambda+2\mu)} M_p.$$

Here,  $M_{\xi}$  and  $M_p$  are mass matrices that correspond to the identity operators in  $Q^h$  and  $W^h$ , respectively. Therefore, one can take the above two as the approximations of  $S_{\xi}$  and  $S_p$ , respectively. In the implementation, one only needs to generate the mass matrices in  $Q^h$  and  $W^h$  and put the proper scalings as shown above. In alignment with the formulations for  $\mathcal{P}_{T_1}$ ,  $\mathcal{P}_{T_2}$ ,  $\mathcal{P}_{T_3}$ , and  $\mathcal{P}_{T_4}$  presented in Eqs. (2.1) and (2.2), the corresponding preconditioners are represented as  $\bar{\mathcal{P}}_{T_1}$ ,  $\bar{\mathcal{P}}_{T_2}$ ,  $\bar{\mathcal{P}}_{T_3}$ , and  $\bar{\mathcal{P}}_{T_4}$ , respectively. This representation is achieved by substituting the exact Schur complements with the aforementioned approximations. Likewise, for the counterparts  $\mathcal{P}_{D_1}$ ,  $\mathcal{P}_{D_2}$ ,  $\mathcal{P}_{D_3}$ , and  $\mathcal{P}_{D_4}$  in Eqs. (2.4) and (2.7), the resulting preconditioners are denoted as  $\bar{\mathcal{P}}_{D_1}$ ,  $\bar{\mathcal{P}}_{D_2}$ ,  $\bar{\mathcal{P}}_{D_3}$ , and  $\bar{\mathcal{P}}_{D_4}$ , respectively.

For practical implementation, for each of  $A_u$ ,  $S_{\xi}$  and  $-S_p$  (because  $S_p$  is negative definite), we apply an incomplete Cholesky factorization, i.e.,

$$\hat{A} = LL^T + R$$

with a drop tolerance. Then  $\hat{A}^{-1}$  is approximated by  $L^{-T}L^{-1}$ . We comment here that other types of inexact solvers can also be applied, for example, the methods introduced in [13, 14, 32]. We denote the corresponding inexact approximations of  $\bar{\mathcal{P}}_{D_1}$  to  $\bar{\mathcal{P}}_{D_4}$  as  $\hat{\mathcal{P}}_{D_1}$  to  $\hat{\mathcal{P}}_{D_4}$ . Similarly, we denote the corresponding inexact approximations of  $\bar{\mathcal{P}}_{T_1}$  to  $\bar{\mathcal{P}}_{T_4}$  as  $\hat{\mathcal{P}}_{T_1}$  to  $\hat{\mathcal{P}}_{T_4}$ . The implementation code, numerical results, and eigenvalues calculations are available at <https://github.com/cmchao2005/Preconditioners-for-Biot-model>.

In our numerical experiments,  $\nu = 0.499$ , and  $\lambda$  and  $\mu$  are computed by using (6.4). All the other physical parameters are set to be 1. The computational

domain is a unit square. The Dirichlet type boundaries  $\Gamma_d$  and  $\Gamma_p$  are the two vertical lines. Other parts of the boundary are Neumann-type. In Table 1, we report the GMRES iteration counts for all preconditioners with the drop tolerance for the incomplete factorizations being  $10^{-3}$ . In Table 2, we summarize the GMRES iteration counts for all preconditioners with the drop tolerance for the incomplete factorizations being  $10^{-4}$ . In all tests, the stopping criterion for the GMRES method is set to be that the  $l^2$  norm of the relative residual is smaller than  $1.0E-6$ . From the numerical experiments, we see that the results based on  $\widehat{\mathcal{P}}_{D_1}^{-1}$  are almost the same as those based on  $\widehat{\mathcal{P}}_{D_2}^{-1}$ , the results based on  $\widehat{\mathcal{P}}_{T_2}^{-1}$  are almost the same as those based on  $\widehat{\mathcal{P}}_{T_3}^{-1}$ . In comparison,  $\widehat{\mathcal{P}}_{D_3}$  outperforms among the block diagonal preconditioners, and  $\widehat{\mathcal{P}}_{T_1}$  outperforms among the block triangular preconditioners. Note that we have theoretically proved that  $\mathcal{P}_{D_3}$  and  $\mathcal{P}_{T_1}$  are positively stable preconditioners. The numerical experiments clearly illustrate the advantages of the positively stable preconditioners. Although we do not rigorously prove that  $\widehat{\mathcal{P}}_{D_3}$  and  $\widehat{\mathcal{P}}_{T_1}$  are positively stable preconditioners in this draft, the corresponding numerical spectral verifications are provided in our code.

Table 1: Comparisons of the different preconditioners. GMRES iteration counts (and CPU time in seconds). The drop tolerance for the incomplete Cholesky factorizations is set to be  $1.0E-3$ .

Block Diagonal Preconditioners				
$h$	$\widehat{\mathcal{P}}_{D_1}^{-1}$	$\widehat{\mathcal{P}}_{D_2}^{-1}$	$\widehat{\mathcal{P}}_{D_3}^{-1}$	$\widehat{\mathcal{P}}_{D_4}^{-1}$
1/16	55 (0.0135)	56 (0.0129)	50 (0.0122)	67 (0.0172)
1/32	108 (0.2671)	108 (0.2605)	96 (0.2213)	127 (0.3559)
1/64	229 (2.645)	229 (2.586)	204 (2.170)	273 (3.538)
1/128	470 (28.94)	470 (29.03)	449 (28.03)	593 (62.61)
1/256	1057 (1027)	1067 (1029)	934 (957.9)	1767 (1765)
Block Triangular Preconditioners				
$h$	$\widehat{\mathcal{P}}_{T_1}^{-1}$	$\widehat{\mathcal{P}}_{T_2}^{-1}$	$\widehat{\mathcal{P}}_{T_3}^{-1}$	$\widehat{\mathcal{P}}_{T_4}^{-1}$
1/16	27 (0.0069)	49 (0.0118)	49 (0.0116)	36 (0.0092)
1/32	55 (0.1004)	97 (0.2237)	97 (0.2187)	72 (0.1350)
1/64	118 (0.936)	206 (2.225)	206 (2.196)	157 (1.432)
1/128	261 (11.00)	434 (42.44)	434 (42.83)	345 (28.44)
1/256	569 (382.1)	947 (980.6)	947 (977.5)	782 (685.1)

Table 2: Comparisons of the different preconditioners. GMRES iteration counts (and CPU time in seconds). The drop tolerance for the incomplete Cholesky factorizations is set to be  $1.0E-4$ .

$h$	Block Diagonal Preconditioners			
	$\widehat{\mathcal{P}}_{D_1}^{-1}$	$\widehat{\mathcal{P}}_{D_2}^{-1}$	$\widehat{\mathcal{P}}_{D_3}^{-1}$	$\widehat{\mathcal{P}}_{D_4}^{-1}$
1/16	36 (0.0127)	36 (0.0114)	32 (0.0118)	41 (0.0139)
1/32	51 (0.3073)	51 (0.2624)	45 (0.2210)	57 (0.3436)
1/64	92 (0.988)	92 (0.947)	80 (0.810)	97 (1.007)
1/128	187 (8.88)	188 (8.96)	167 (7.83)	205 (10.54)
1/256	373 (89.18)	373 (86.59)	353 (82.71)	455 (119.4)
Block Triangular Preconditioners				
$h$	$\widehat{\mathcal{P}}_{T_1}^{-1}$	$\widehat{\mathcal{P}}_{T_2}^{-1}$	$\widehat{\mathcal{P}}_{T_3}^{-1}$	$\widehat{\mathcal{P}}_{T_4}^{-1}$
	15 (0.0073)	31 (0.0111)	31 (0.0110)	19 (0.0085)
1/32	24 (0.0961)	45 (0.2188)	45 (0.2415)	29 (0.1259)
1/64	45 (0.414)	84 (0.854)	84 (0.841)	54 (0.502)
1/128	97 (4.31)	170 (7.81)	170 (7.88)	119 (5.05)
1/256	206 (39.28)	350 (80.22)	350 (79.07)	267 (54.99)

## 7 Concluding remarks

In this paper, both nested Schur complement and additive Schur complement based preconditioners are constructed for the twofold and block tridiagonal linear systems. The polynomial equations of the preconditioned matrices are analyzed. It is shown that by properly selecting the sign in front of each Schur complement, some preconditioners are positively stable. Numerical experiments based on the Biot model are provided to show that positively stable preconditioners outperform other preconditioners. More clearly, when inexact elliptic approximations are incorporated, the inexact versions of the positively stable preconditioners are more efficient.

For block tridiagonal systems, by comparing the theoretical analysis for the nested Schur complement based preconditioners and that for the additive type preconditioners, our argument is that permutation is important and necessary when designing preconditioners. These results are instructive for devising the corresponding inexact versions of the preconditioners and iterative methods [4].

## Appendix A. An alternative analysis of the Hurwitz stability of polynomials (3.5)

A multivariate  $f \in \mathbb{C}[z_1, \dots, z_n]$  is said to be weakly Hurwitz stable if  $f$  is either identically zero or nonvanishing whenever  $\Re(z_i) > 0$  for any  $i \in [n]$ , where  $\Re(z)$  is the real part of  $z$  for  $z \in \mathbb{C}$ . Denote by  $\{r_i\}$  and  $\{s_j\}$  the all real zeros of  $f, g \in \mathbb{R}[x]$ , respectively. We say that  $g$  alternates left of  $f$  if  $\deg(f) = \deg(g) = n$  and

$$s_n \leq r_n \leq \dots \leq s_2 \leq r_2 \leq s_1 \leq r_1, \quad (\text{A.1})$$

and that  $g$  interlaces  $f$  if  $\deg(f) = \deg(g) + 1 = n$  and

$$r_n \leq s_{n-1} \leq \dots \leq s_2 \leq r_2 \leq s_1 \leq r_1. \quad (\text{A.2})$$

Let  $g \preceq f$  denote either  $g$  alternates left  $f$  or  $g$  interlaces of  $f$ . Here, we denote  $g \ll f$  if  $g \preceq f$  (resp.,  $f \preceq g$ ) and the leading coefficients of  $f, g$  have the same (resp., opposite) sign.

Assume that a polynomial  $f(x) = \sum_{k=0}^n f_k x^k$  has degree  $n$ , then let

$$f^E(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} f_{2k} x^k \quad \text{and} \quad f^O(x) = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} f_{2k+1} x^k.$$

**Theorem A.1** (Hermite-Biehler Theorem [6,35, Theorem 6.3.4]). *A polynomial  $f(x)$  with real coefficients is weakly Hurwitz stable if and only if  $f^E(x)$  and  $f^O(x)$  have only real and nonpositive zeros, and  $f^O(x) \ll f^E(x)$ .*

**Theorem A.2.** *Let  $p_n(x)$  satisfy the recurrence relation (3.5) with  $p_0(x) = 1$ ,  $p_1(x) = x - 1$ . Then, for any  $n \geq 0$ , all zeros of polynomial  $p_n(x)$  lie in the right half-plane.*

*Proof.* Obviously, all zeros of  $p_n(x)$  lie in the right half-plane if and only if all zeros of  $p_n(-x)$  lie in the left closed half-plane, i.e., weakly Hurwitz stable. Let

$$f_n(x) = \sum_{k=0}^n f_k x^k := (-1)^n p_n(-x).$$

Hence,  $f_n(x)$  satisfies recurrence relation

$$f_n(x) = x f_{n-1}(x) + f_{n-2}(x), \quad (\text{A.3})$$

where  $f_0(x) = 1$  and  $f_1(x) = x + 1$ . Note that  $f_n(x)$  has only nonnegative coefficients for all  $n \geq 0$ .

Let

$$f_n^E(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} f_{2k} x^k \quad \text{and} \quad f_n^O(x) = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} f_{2k+1} x^k.$$

Due to Theorem A.1, it suffices to show that  $f_n^O(x) \ll f_n^E(x)$  for any  $n \geq 0$ . By the recurrence (A.3), we get that

$$f_n^E(x) = f_{n-2}^E(x) + x f_{n-1}^O(x), \quad (\text{A.4})$$

$$f_n^O(x) = f_{n-1}^E(x) + f_{n-2}^O(x). \quad (\text{A.5})$$

By induction on  $n$ , it is easy to see that

$$f_{2n}^E(x) = f_{2n+1}^E(x) \quad \text{and} \quad f_{2n-1}^O(x) = f_{2n}^O(x). \quad (\text{A.6})$$

Subsequently, we will prove the following results through an induction argument based on  $n$ :

- (1)  $f_{n-2}^E(x) \ll f_n^E(x)$ ,
- (2)  $f_{n-2}^O(x) \ll f_n^O(x)$ ,
- (3)  $f_{n-1}^O(x) \ll f_{n-2}^E(x)$ ,
- (4)  $f_{n-1}^O(x) \ll f_n^E(x)$ .

Firstly,  $f_n^O(x) \ll f_n^E(x)$  follows from interlacing relations (3) and (4) by (A.6). When  $n=4$ , by Table 3 it is routine to verify that (1)-(4) are true by Table 4. Assume that (1)-(4) are true for an  $n > 4$ . We need to prove that

- (a)  $f_{n-1}^E(x) \ll f_{n+1}^E(x)$ ,
- (b)  $f_{n-1}^O(x) \ll f_{n+1}^O(x)$ ,
- (c)  $f_n^O(x) \ll f_{n-1}^E(x)$ ,
- (d)  $f_n^O(x) \ll f_{n+1}^E(x)$ .

By (A.5) and the interlacing relations

$$f_n^O(x) \ll f_{n-1}^E(x) + f_{n-2}^O(x) \quad \text{and} \quad f_n^O(x) \ll -f_{n-2}^O(x),$$

Table 3: The terms of  $p_n(x)$  for  $1 \leq n \leq 4$ .

$n$	$p_n(x)$
1	$x-1$
2	$x^2-x+1$
3	$x^3-x^2+2x-1$
4	$x^4-x^3+3x^2-2x+1$

Table 4: The terms of  $f_n^O(x)$  and  $f_n^E(x)$  for  $2 \leq n \leq 4$ .

$n$	$f_n^O(x)$	$f_n^E(x)$
2	1	$x+1$
3	$x+2$	$x+1$
4	$x+2$	$x^2+3x+1$

we have  $f_n^O(x) \ll f_{n-1}^E(x)$ , i.e., (c) holds. Moreover, by (A.4) and combining

$$f_{n-1}^E(x) \ll x f_n^O(x) \quad \text{and} \quad f_{n-1}^E(x) \ll f_{n-1}^E(x)$$

we derive that  $f_{n-1}^E(x) \ll f_{n+1}^E(x)$ , i.e., (a) holds. Note that  $f_{n+1}^O(x) = f_n^E(x) + f_{n-1}^O(x)$ , hence (b) follows from (4) and  $f_{n-1}^O(x) \ll f_{n-1}^O(x)$ . Similarly, due to  $f_{n+1}^E(x) = f_{n-1}^E(x) + x f_n^O(x)$  by (A.4), (d) is immediately derived by (c) and  $f_n^O(x) \ll x f_n^O(x)$ . This completes the proof.  $\square$

## Acknowledgments

The authors would like to thank Mingjian Ding, and Baoxuan Zhu for providing an alternative proof of the Hurwitz stability of polynomials (3.5). They also thank Jarle Sogn for communicating on Schur complement based preconditioners.

The work of M. Cai is partially supported by the NIH-RCMI (Grant No. 347U54MD013376), by the affiliated project award from the Center for Equitable Artificial Intelligence and Machine Learning Systems at Morgan State University (Project ID 02232301), and by the National Science Foundation awards (Grant No. 1831950). The work of G. Ju is supported in part by the National Key R&D Program of China (Grant No. 2017YFB1001604). The work of J. Li is partially supported by the National Natural Science Foundation of China (Grant No. 11971221) and by the Shenzhen Sci-Tech Fund (Grant Nos. RCJC20200714114556020, JCYJ20170818153840322, JCYJ20190809150413261), and by the Guangdong Provincial Key Laboratory of Computational Science and Material Design (Grant No. 2019B030301001).

## References

- [1] J. Adler, F. Gaspar, X. Hu, C. Rodrigo, and L. Zikatanov, *Robust block preconditioners for Biot's model*, in: Domain Decomposition Methods in Science and Engineering XXIV, Lecture Notes in Computational Science and Engineering, Springer, Cham 125 (2018), 3–16.
- [2] S. Barnett and D. Šiljak, *Routh's algorithm: A centennial survey*, SIAM Rev. 19(3) (1977), 472–489.
- [3] F. Beik and M. Benzi, *Block preconditioners for saddle point systems arising from liquid crystal directors modeling*, Calcolo 55(3) (2018), 1–16.
- [4] F. Beik and M. Benzi, *Iterative methods for double saddle point systems*, SIAM J. Matrix Anal. Appl. 39(2) (2018), 902–921.

- [5] M. Benzi, G. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numer. 14 (2005), 1–137.
- [6] J. Borcea and P. Brändén, *Multivariate Polya-Schur classification problems in the Weyl algebra*, Proc. Lond. Math. Soc. 101(3) (2010), 73–104.
- [7] S. Bradley and C. Greif, *Eigenvalue bounds for double saddle-point systems*, IMA J. Numer. Anal. 43(6) (2023), 3564–3592.
- [8] M. Cai, *Analysis of some projection method based preconditioners for models of incompressible flow*, Appl. Numer. Math. 90 (2015), 77–90.
- [9] M. Cai, G. Ju, and J. Li, *Overlapping Schwarz preconditioners for the Biot model*, Preprint, 2023.
- [10] M. Cai, M. Mu, and J. Xu, *Preconditioning techniques for a mixed Stokes/Darcy model in porous media applications*, J. Comput. Appl. Math. 233 (2009), 346–355.
- [11] M. Cai, A. Nonaka, J. Bell, B. Griffith, and A. Donev, *Efficient variable-coefficient finite-volume Stokes solvers*, Comm. Comput. Phys. 16(5) (2014), 1263–1297.
- [12] Y. Cao, *Shift-splitting preconditioners for a class of block three-by-three saddle point problems*, Appl. Math. Lett. 96 (2019), 40–46.
- [13] Y. Cao and M. Neytcheva, *Cell-by-cell approximate Schur complement technique in preconditioning of meshfree discretized piezoelectric equations*, Numer. Linear Algebra Appl. 28 (2021), e2362.
- [14] Y. Cao, Q.-Q. Shen, and Y.-T. Chen, *Additive inexact block triangular preconditioners for saddle point problems arising in meshfree discretization of piezoelectric equations*, East Asian J. App. Math. 12 (2022), 381–405.
- [15] Z. Cao, *A class of constraint preconditioners for nonsymmetric saddle point matrices*, Numer. Math. 103(1) (2006), 47–61.
- [16] Z. Cao, *Positive stable block triangular preconditioners for symmetric saddle point problems*, Appl. Numer. Math. 7 (2007), 899–910.
- [17] B. Cockburn and K. Shi, *Devising HDG methods for Stokes flow: An overview*, Comput. Fluids. 98 (2014), 221–229.
- [18] M. Eiermann and O. Ernst, *Geometric aspects of the theory of Krylov subspace methods*, Acta Numer. 10 (2001), 251–312.
- [19] B. Fischer, A. Ramage, D. Silvester, and A. Wathen, *Minimum residual methods for augmented systems*, BIT Numer. Math. 38(3) (1998), 527–543.
- [20] F. Gantmacher, *Applications of the Theory of Matrices*, Interscience, 1959.
- [21] G. Gatica, R. Oyarzúa, and F. Sayas, *A twofold saddle point approach for the coupling of fluid flow with nonlinear porous media flow*, IMA J. Numer. Anal. 32 (2012), 845–887.
- [22] Q. Hong, J. Kraus, M. Lymbery, F. Philo, *Parameter-robust Uzawa-type iterative methods for double saddle point problems arising in Biot's consolidation and multiple-network poroelasticity models*, Math. Models Methods Appl. Sci. 30(13) (2020), 2523–2555.
- [23] J. Howell and N. Walkington, *Inf-sup conditions for twofold saddle point problems*, Numer. Math. 118 (2011), 663–693.

- [24] N. Huang and C. Ma, *Spectral analysis of the preconditioned system for the  $3 \times 3$  block saddle point problem*, Numer. Algorithms 81(2) (2019), 421–444.
- [25] I. Ipsen, *A note on preconditioning nonsymmetric matrices*, SIAM J. Sci. Comput. 23 (2001), 1050–1051.
- [26] J. Lee, K. Mardal, and R. Winther, *Parameter-robust discretization and preconditioning of Biot's consolidation model*, SIAM J. Sci. Comput. 39 (2017), A1–A24.
- [27] Y. Lin and Y. Wei, *A note on constraint preconditioners for nonsymmetric saddle point problems*, Numer. Linear Algebra Appl. 14(8) (2007), 659–664.
- [28] K. Mardal, B. Nielsen, and M. Nodaas, *Robust preconditioners for PDE-constrained optimization with limited observations*, BIT Numer. Math. 57(2) (2017), 405–431.
- [29] K. Mardal and R. Winther, *Preconditioning discretizations of systems of partial differential equations*, Numer. Lin. Alg. Appl. 18(1) (2011), 1–40.
- [30] T. Mathew, *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*, Lecture Notes in Computational Science and Engineering, Vol. 61, Springer, 2008.
- [31] M. Murphy, G. Golub, and A. Wathen, *A note on preconditioning for indefinite linear systems*, SIAM J. Sci. Comput. 21 (2000), 1969–1972.
- [32] M. Neytcheva, *On element-by-element Schur complement approximations*, Linear Algebra Appl. 434 (2011), 2308–2324.
- [33] R. Oyarzúa and R. Ruiz-Baier, *Locking-free finite element methods for poroelasticity*, SIAM J. Numer. Anal. 54(5) (2016), 2951–2973.
- [34] J. Pearson and A. Wathen, *A new approximation of the Schur complement in preconditioners for PDE-constrained optimization*, Numer. Linear Alg. Appl. 19 (2012), 816–829.
- [35] Q. I. Rahman and G. Schmeisser, *Analytic Theory of Polynomials*, London Math. Soc. Monographs New Series 26, Oxford University Press, 2002.
- [36] A. Ramage and J. Gartland, *A preconditioned nullspace method for liquid crystal director modeling*, SIAM J. Sci. Comput. 35 (2013), B226–B247.
- [37] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, 2003.
- [38] J. Sogn and W. Zulehner, *Schur complement preconditioners for multiple saddle point problems of block tridiagonal form with application to optimization problems*, IMA J. Numer. Anal. 39 (2019), 1328–1359.
- [39] A. Toselli and O. Widlund, *Domain Decomposition Methods-Algorithms and Theory*, Springer Science & Business Media, 2006.
- [40] C. Wang and P. Sun, *An augmented Lagrangian Uzawa iterative method for solving double saddle-point systems with semidefinite  $(2,2)$  block and its application to DLM/FD method for elliptic interface problems*, Commun. Comput. Phys. 30(1) (2021), 124–143.
- [41] X. Xie and H.-B. Li, *A note on preconditioning for the  $3 \times 3$  block saddle point problem*, Comput. Math. Appl. 79 (2020), 3289–3296.