DELFT UNIVERSITY OF TECHNOLOGY

REPORT 15-01

Convergence Analysis of Multilevel Sequentially Semiseparable Preconditioners

Yue Qiu, Martin B. van Gijzen, Jan-Willem van Wingerden
Michel Verhaegen, and Cornelis Vuik

ISSN 1389-6520

Reports of the Delft Institute of Applied Mathematics

Delft 2015

CONVERGENCE ANALYSIS OF MULTILEVEL SEQUENTIALLY SEMISEPARABLE PRECONDITIONERS*

YUE QIU†, MARTIN B. VAN GIJZEN‡, JAN-WILLEM VAN WINGERDEN†, MICHEL VERHAEGEN†, AND CORNELIS VUIK‡

Abstract. Multilevel sequentially semiseparable (MSSS) matrices form a class of structured matrices that have low-rank off-diagonal structure, which allows the matrix-matrix operations to be performed in linear computational complexity. MSSS preconditioners are computed by replacing the Schur complements in the block LU factorization of the global linear system by MSSS matrix approximations with low off-diagonal rank. In this manuscript, we analyze the convergence properties of such preconditioners. We show that the spectrum of the preconditioned system is contained in a circle centered at $(1, 0)$ and give an analytic bound of the radius of this circle. This radius can be made arbitrarily small by properly setting a parameter in the MSSS preconditioner. Our results apply to a wide class of linear systems. The system matrix can be either symmetric or unsymmetric, definite or indefinite. We demonstrate our analysis by numerical experiments.

Key words. multilevel sequentially semiseparable preconditioners, convergence analysis, saddle-point systems, Helmholtz equation

AMS subject classifications. 15B99, 65Fxx, 93C20, 65Y20

1. Introduction. The most time consuming part for many numerical simulations in science and engineering is to solve one or more linear systems of the following type

$$Kx = b, \tag{1.1}$$

where $K = [K_{ij}]$ is an $n \times n$ matrix and b is a given right-hand-side vector of compatible size. For the discretization of partial differential equations (PDEs), the system matrix K is usually large and sparse. Many efforts have been dedicated to finding efficient numerical solution for such systems. Krylov subspace methods, such as the conjugate gradient (`CG`), minimal residual (`MINRES`), generalized minimal residual (`GMRES`) and induced dimension reduction (`IDR(s)`) methods [18, 26, 36, 39], have attracted considerable attention over the last decades. When Krylov subspace methods are applied to solve large linear systems, preconditioning is necessary to improve the robustness and convergence of such iterative solvers. This manuscript studies the multilevel sequentially semiseparable (MSSS) preconditioners. The efficiency of MSSS preconditioners has been shown in [31] for difficult linear systems arising from PDE-constrained optimization problems. Later, it is extended to solve the general computational fluid dynamics (CFD) problems in [32]. The computational results given by [31, 32] demonstrate that the induced dimension reduction (`IDR(s)`) method can compute the solution in 2-4 iterations by using the MSSS preconditioner for saddle-point system of the following type

$$\underbrace{\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix}}_{\mathcal{A}} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}. \tag{1.2}$$

*This research is partly supported by the NWO Veni Grant # 11930 "Reconfigurable Floating Wind Farms".

†Delft Center for Systems and Control, Delft University of Technology, 2628 CD, Delft, the Netherlands. (Y.Qiu@tudelft.nl, Y.Qiu@gmx.com).

‡Delft Institute of Applied Mathematics, Delft University of Technology, 2628 CD, Delft, the Netherlands. (M.B.vanGijzen@tudelft.nl).

Here, $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite, $B \in \mathbb{R}^{m\times n}$ has full rank, $C \in \mathbb{R}^{m\times m}$ is symmetric positive semidefinite and usually $m \leq n$. Such systems arise in PDE-constrained optimization problems [42, 33], computational fluid fluid dynamics [46, 15], optimal flow control [34, 27] et al., cf. [3] for a general survey of the saddle-point systems and numerical solutions.

It is shown in [31, 32] that all the sub-matrices of the saddle-point systems (1.2) have a multilevel sequentially semiseparable (MSSS) structure. This sub-structure can be exploited to obtain a global MSSS structure of the form (1.1) by performing a simple permutation. This permutation unites the linear systems from the discretization of scalar PDEs such as the Helmholtz equation together with coupled PDEs, e.g. the Stokes equation and gives a general linear system of the form (1.1) with a global MSSS structure. In this way, we can obtain a global factorization of the system by using MSSS matrix computations in linear computation complexity. The global factorization can be computed with a prescribed accuracy, which gives a flexible and numerically efficient way to solve a wide class of linear systems. The central concept for this factorization is that the off-diagonal blocks of the Schur complements are observed to have numerical low-rank [9, 2]. The low numerical off-diagonal blocks can be approximated efficiently and accurately by structured matrices, such as the multilevel sequentially semiseparable (MSSS) matrices, hierarchically semiseparable (HSS) matrices, hierarchical matrices, et al.

Multilevel sequentially semiseparable (MSSS) matrix generalize the sequentially semiseparable (SSS) matrices [8] to the multidimensional case. They can be directly inferred from interconnected systems [35]. Early study of preconditioning by using SSS matrix computations for symmetric positive definite systems can be found in [19] while MSSS matrix computations have been extensively studied in [28, 31, 32] for preconditioning unsymmetric and saddle-point systems. The advantage of MSSS matrix computations is their simplicity and low computational cost, which is $\mathcal{O}(r_k^3 N)$. Here, N is the number of blocks, r_k is the rank of the off-diagonal blocks and is usually much smaller compared with N [8, 45]. Related structured matrices include the hierarchical matrices ($\mathcal{H}$-matrix) [20], $\mathcal{H}^2$-matrices [21, 5], hierarchically semiseparable (HSS) matrices [7, 48], with computational complexity $\mathcal{O}(N \log^\alpha N)$ for some moderate α. Here N is the size of the matrix. $\mathcal{H}$-matrices originate from approximating the kernel of integral equations and have been extended to elliptic PDEs [2, 22]. $\mathcal{H}^2$-matrices and HSS matrices are specific subsets of $\mathcal{H}$-matrices and HSS matrices have been successfully applied in the multi-frontal solvers [48]. Some recent efforts have been devoted to preconditioning symmetric positive definite systems by exploiting the HSS matrix structure [47] and unsymmetric systems by $\mathcal{H}$-matrix computations [22, 6]. To keep a clear structure of this manuscript, we only focus on the MSSS preconditioning techniques.

In this manuscript, we present a full convergence analysis of the MSSS preconditioners for a wide class of linear systems. The system matrix can be either symmetric or unsymmetric, definite or indefinite, where saddle-point systems, discretized Helmholtz equations, and discretized convection-diffusion equations are automatically covered. Our analysis gives an analytic bound for the spectrum of the preconditioned system. We show that the spectrum of the preconditioned system is contained in a circle centered at $(1, 0)$ and give an analytic bound for the radius of this circle. This radius can be made arbitrarily small by properly setting a parameter in the MSSS preconditioner.

Some related work includes [1] and [24]. Both analyses apply only to symmetric

positive definite systems. The analysis for MSSS preconditioners in [24] is restricted to 1-level MSSS matrix computations, while our analysis can be applied to 2-level MSSS matrix computations. Our work in this manuscript is closely related to [1]. Our contributions include: (1) We extend the work in [1, 24] from the symmetric positive definite case to the general linear systems; (2) Our analysis can also be applied to saddle-point systems that are 2×2 block systems, while the analysis in [1, 24] only applies to symmetric positive definite linear systems that arise from discretization of scalar PDEs; (3) We give an analytic bound for the error introduced by the model order reduction that is necessary for the MSSS preconditioning technique, which has not been studied before; (4) The analysis for MSSS preconditioning in [24] only concerns 1-level MSSS matrix computations, while our analysis also includes the 2-level MSSS cases; (5) For the first time, we apply this MSSS preconditioning technique to the Helmholtz equation.

The structure of our manuscript is as follows. We give a brief introduction of the MSSS preconditioning technique in Section 2, and analyze its convergence in Section 3. Section 4 studies the numerical stability of this preconditioning technique and gives a sufficient condition to avoid breakdown. The underlying condition can be satisfied by properly setting a parameter. We show how to choose this parameter in Section 5. Numerical experiments are given in Section 6, while conclusions are drawn in the final part. A companion technical report [30] is available online and contains more numerical results to illustrate our analysis for the convergence of MSSS preconditioners.

2. Multilevel Sequentially Semiseparable Preconditioners. To start with, we first introduce the sequentially semiseparable matrices and the generators definition for such matrices is given by Definition 2.1.

Definition 2.1 ([8]). *Let A be an $N \times N$ matrix with SSS matrix structure and let n positive integers $m_1,\ m_2,\ \cdots\ m_n$ satisfy $N = m_1 + m_2 + \cdots + m_n$ such that A can be written in the following block-partitioned form*

$$A_{ij} = \begin{cases} U_i W_{i+1} \cdots W_{j-1} V_j, & \text{if} \quad i < j; \\ D_i, & \text{if} \quad i = j; \\ P_i R_{i-1} \cdots R_{j+1} Q_j, & \text{if} \quad i > j. \end{cases} \tag{2.1}$$

The matrices U_i, W_i, V_i, D_i, P_i, R_i, Q_i are matrices whose sizes are compatible for matrix-matrix product when their sizes are not mentioned. They are called generators of the SSS matrix A.

Basic operations such as addition, multiplication and inversion are closed under the SSS matrix structure and can be performed in linear computational complexity. Multilevel sequentially semiseparable matrices generalize the SSS matrices to the multi-dimensional case. Similar to Definition 2.1 for SSS matrices, the generators representation for MSSS matrices, specifically the k-level SSS matrices, is defined in Definition 2.2.

Definition 2.2. *The matrix A is said to be a k-level SSS matrix if all its generators are $(k-1)$-level SSS matrices. The 1-level SSS matrix is the SSS matrix that satisfies Definition 2.1.*

The MSSS matrix structure can be inferred directly from the discretization of PDEs, which is studied in [28, 31, 32]. With these definitions, we start to introduce the MSSS preconditioners for discretized partial differential equations (PDEs). Consider the following PDE

$$\mathcal{L}u = f, \text{ with } u = u_D \text{ on } \Gamma_D,$$

on a square domain $\Omega \in \mathbb{R}^d$ with $d = 2,$ or 3, $\mathcal{L}$ is a linear differential operator, and $\Gamma_D = \partial\Omega$. Discretizing the PDE using the finite difference method or the finite element method and using lexicographical to order the grid points gives the following linear system,

$$Kx = b,$$

where the stiffness matrix K is block tridiagonal of the following type

$$K = \begin{bmatrix} K_{1,1} & K_{1,2} & & & \\ K_{2,1} & K_{2,2} & K_{2,3} & & \\ & K_{3,2} & K_{3,3} & \ddots & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & K_{N,N} \end{bmatrix}. \tag{2.2}$$

Here $K_{i,j}$ is again a tridiagonal matrix for $d = 2$ and block tridiagonal matrix for $d = 3$.

For discretized scalar PDEs using uniform mesh, it is quite natural to infer that the stiffness matrix K has an MSSS matrix structure, i.e., a 2-level MSSS structure for $d = 2$ and a 3-level MSSS structure for $d = 3$. Discretized coupled PDEs, such as the Stokes equation, and linearized Navier-Stoke equation yield a saddle-point system of the form (1.2). It is shown in [32], that all the sub-matrix in (1.2) have an MSSS structure and can be permuted into a global MSSS structure that has the same form as (2.2) with $K_{i,j}$ a 2-level SSS matrix for $d = 3$ and 1-level SSS matrix for $d = 2$.

For a strongly regular $N \times N$ block matrix K, it admits the block factorization that is given by $K = LSU$. Here we say that a matrix is strongly regular if all the leading principle sub-matrices are nonsingulari. S is a block diagonal matrix with its i-th diagonal block given by

$$S_i = \begin{cases} K_{i,i} & \text{if } i = 1 \\ K_{i,i} - K_{i,i-1}S_{i-1}^{-1}K_{i-1,i} & \text{if } 2 \le i \le N \end{cases}, \tag{2.3}$$

where S_i is the Schur complement at the i-th step. The matrix L and U are block bidiagonal matrix of the lower-triangular form and upper-triangular form, respectively. They are obtained by computing

$$L_{i,j} = \begin{cases} I & \text{if } i = j \\ K_{i,j}S_j^{-1} & \text{if } i = j+1 \end{cases}, \text{ and } U_{i,j} = \begin{cases} I & \text{if } j = i \\ S_i^{-1}K_{i,j} & \text{if } j = i+1 \end{cases}.$$

To compute such a factorization, one needs to compute the Schur complements via (2.3). This is computationally expensive both in time and memory since the Schur complement S_i is a full matrix. Some earlier papers [10, 23] propose for symmetric positive definite systems to approximate S_i by using the off-diagonal decay property of the inverse of a symmetric positive definite tridiagonal matrix. Alternatively, an incompletely factorization can be made to reduce the fill-in within the bandwidth for such a factorization [36]. However, these methods do not yield satisfactory performance. In [2, 9], it is stated that the Schur complement S_i from the factorization of a discretized symmetric positive definite differential operator has low off-diagonal rank and can be approximated by an $\mathcal{H}$-matrix or SSS matrix. In this manuscript,

we use SSS matrices to approximate the Schur complements in the above factorization for a general class of linear systems. Note that the $\mathcal{H}$-matrix computations have mainly been applied to approximate the Schur complements that are either symmetric positive definite [2, 1] or unsymmetric, whose eigenvalues have positive real part [22]. Some recent efforts have been made to solve the symmetric indefinite Helmholtz equation [16].

For the approximated Schur complement by MSSS matrix computations denoted by $\tilde{S}_i, (i = 1,\ 2,\ \cdots,\ N)$, we have the MSSS preconditioner $\tilde{K}$ that is given by

$$\tilde{K} = \tilde{L}\tilde{S}\tilde{U}. \tag{2.4}$$

Here

$$\tilde{L}_{i,j} = \begin{cases} I & \text{if } i = j \\ K_{i,j}\tilde{S}_j^{-1} & \text{if } i = j+1 \end{cases}, \quad \tilde{U}_{i,j} = \begin{cases} I & \text{if } j = i \\ \tilde{S}_i^{-1}K_{i,j} & \text{if } j = i+1 \end{cases},$$

and $\tilde{S}$ is a block-diagonal matrix with $\tilde{S}_i, (i = 1,\ 2,\ \cdots,\ N)$ as its diagonal blocks.

The Schur complement always corresponds to a problem that is one dimension lower than the linear system, i.e., for a 2D system, the Schur complement is of 1D, and 2D for a 3D system. When applying MSSS preconditioners to 3D problems, one needs 2-level MSSS matrix computations to approximate the Schur complement. How to reduce the off-diagonal rank of a 2-level MSSS matrix efficiently is still an open problem [9, 11, 31]. This makes extending MSSS preconditioning technique from 2D to 3D nontrival, and some extra efforts need to be devoted, cf. [11, 31]. To keep consistent of this manuscript, we only focus on the convergence analysis of MSSS preconditioner for 2D systems.

3. Convergence Analysis. In this section, we analyze the convergence of the MSSS preconditioner. Some recent work devoted to the analysis of structured preconditioners are in [1, 24]. In [24], the nested dissection method is used to order the unknowns of the discretized diffusion-reaction equation in 2D and the symmetric positive definite Schur complements are approximated by SSS matrix and HSS matrix computations, respectively. Analytic bounds of the spectrum of the preconditioned system are given. In [1], $\mathcal{H}$-matrix computations are applied to preconditioning the 2D symmetric positive definite Helmholtz equation. Both studies focus on the symmetric positive definite case.

In [1], it is stated that the key point for the $\mathcal{H}$-matrix preconditioner for symmetric positive definite systems is not how well the approximate Schur complement denoted by $\tilde{S}_i$ approximates the exact Schur complement S_i, but how small the distance between $\tilde{S}_i$ and $K_{i,i} - K_{i,i-1}\tilde{S}_{i-1}^{-1}K_{i-1,i}$ is. This statement is denoted by the so-called "condition ε" in [1]. In this manuscript, we also make use of this condition for convergence analysis, which is given by the following definition.

Definition 3.1 (Condition ε [1]). *There exists a constant ε such that*

$$\begin{aligned} \left\|\tilde{S}_1 - K_{1,1}\right\|_2 &\le \varepsilon, \\ \left\|\tilde{S}_i - \left(K_{i,i} - K_{i,i-1}\tilde{S}_{i-1}^{-1}K_{i-1,i}\right)\right\|_2 &\le \varepsilon, \end{aligned} \tag{3.1}$$

hold for $2 \le i \le N$.

If condition ε in Definition 3.1 holds, we have the following lemma that gives the distance between the preconditioner and the original system.

LEMMA 3.2. *Let K be a nonsingular matrix that has the form of* (2.2)*, suppose $\tilde{S}_i,\ i = 1, 2, \ldots, N$ in* (3.1) *are nonsingular and condition ε holds. The MSSS preconditioner is given by* (2.4)*, and*

$$\left\| K - \tilde{K} \right\|_2 \leq \varepsilon.$$

Proof. For the proof of this lemma, cf. [1]. ∎

Next, we introduce how to compute an MSSS preconditioner that satisfies Lemma 3.2. Here, we assume the exact arithmetic.

LEMMA 3.3. *Let the lower semiseparable order and upper semiseparable order are defined as the maximal rank of the lower off-diagonal blocks and upper off-diagonal blocks, respectively. For a nonsingular SSS matrix A with lower semiseparable order and upper semiseparable order r_l and r_u, the exact inverse of A can be computed using SSS matrix computations in linear computational complexity provided that r_l and r_u are much smaller than the size of A. The inverse of A is again an SSS matrix with r_l and r_u as its lower and upper semiseparable order, respectively.*

Proof. This can be shown by carefully checking the arithmetic for inverting SSS matrices described in [8, 13]. ∎

LEMMA 3.4. *Let SSS matrices A and B are with compatible sizes and properly partitioned blocks for matrix-matrix addition and multiplication, then $A + B$ and AB can be computed exactly using SSS matrix computations in linear computational complexity if no model order reduction is performed.*

Proof. Proof of this lemma is given by checking the algorithms for SSS matrices introduced in [8, 14]. ∎

For the 2D matrix K of the form in (2.2), all its sub-blocks are SSS matrices, therefore we have the following corollary that shows how the condition ε in Definition 3.1 can be satisfied.

COROLLARY 3.5. *Suppose $\tilde{S}_i,\ i = 1, 2, \ldots, N$ in* (3.1) *are nonsingular, then the condition ε can be satisfied by applying the following procedure.*

1. *Invert $\tilde{S}_{i-1}$ using the SSS inversion algorithm.*
2. *Compute $K_{i,i} - K_{i,i-1}\tilde{S}_{i-1}^{-1}K_{i-1,i}$ using SSS matrix computations without performing the model order reduction.*
3. *Perform the model order reduction for $K_{i,i} - K_{i,i-1}\tilde{S}_{i-1}^{-1}K_{i-1,i}$ by choosing a proper semiseparable order r_k or a proper bound τ for the discarded singular values, which will be introduced in Section 5, such that the condition*

$$\left\| \tilde{S}_i - \left(K_{i,i} - K_{i,i-1}\tilde{S}_{i-1}^{-1}K_{i-1,i} \right) \right\|_2 \leq \varepsilon$$

is satisfied.

Proof. According to Lemma 3.3 and Lemma 3.4, both step 1 and 2 can be performed exactly. By applying the Hankel blocks approximation introduced in [8, 31], $K_{i,i} - K_{i,i-1}\tilde{S}_{i-1}^{-1}K_{i-1,i}$ can be approximated by an SSS matrix $\tilde{S}_i$ in a prescribed accuracy ε measured in the matrix 2-norm by choosing a properly set parameter in the Hankel block approximations. The details will be introduced in Section 5. ∎

REMARK 3.1. To satisfy condition ε, we need to apply Corollary 3.5 to compute the MSSS preconditioner. This is computationally cheaper and feasible, because the semiseparable order of $K_{i,i}$, $\tilde{S}_{i-1}$, $K_{i,i-1}$ and $K_{i-1,i}$ are small. Performing step 2 in Corollary 3.5 just increases the semiseparable order slightly. Here, the semiseparable order is defined as the maximum off-diagonal rank. For details of the increase of the

semiseparable order, cf. [14]. After performing the model order reduction in step 3, the semiseparable order is reduced and bounded. This gives the approximate Schur complements $\tilde{S}_i$ with small semiseparable order.

Lemma 3.2 gives the distance between the preconditioner and the original system matrix while Corollary 3.5 illustrates how to satisfy the condition ε. Normally, we do not consider this distance, but the distance between the preconditioned matrix and the identity matrix. Next, we give an analytic bound of this distance.

THEOREM 3.6. *Let a nonsingular matrix K be of the form* (2.2) *and let the condition ε in Definition 3.1 hold by Corollary 3.5. If $\tilde{S}_i$ $(i = 1, 2, \ldots, N)$ are nonsingular and $\varepsilon < \varepsilon_0$, then the MSSS preconditioner is given by* (2.4) *and*

$$\left\| I - \tilde{K}^{-1} K \right\|_2 \leq \frac{\varepsilon}{\varepsilon_0 - \varepsilon}.$$

Here ε_0 is the smallest singular value of K.

Proof. Since $\tilde{S}_i$ $(i = 1, 2, \ldots, N)$ are nonsingular, $\tilde{K}$ is nonsingular. Then,

$$\left\| I - K^{-1}\tilde{K} \right\|_2 = \left\| K^{-1}(K - \tilde{K}) \right\|_2 \leq \left\| K^{-1} \right\|_2 \left\| K - \tilde{K} \right\|_2 \leq \varepsilon \left\| K^{-1} \right\|_2 = \frac{\varepsilon}{\varepsilon_0}.$$

Since $\varepsilon < \varepsilon_0$, we have $\dfrac{\varepsilon}{\varepsilon_0} < 1$, then the Neumann series

$$I + \left(I - K^{-1}\tilde{K} \right) + \left(I - K^{-1}\tilde{K} \right)^2 + \cdots\cdots$$

converges to $\left(I - (I - K^{-1}\tilde{K}) \right)^{-1} = \tilde{K}^{-1}K$. This in turn gives,

$$\begin{aligned}
\left\| \tilde{K}^{-1} K \right\|_2 &= \left\| I + \left(I - K^{-1}\tilde{K} \right) + \left(I - K^{-1}\tilde{K} \right)^2 + \cdots \right\|_2 \\
&\leq 1 + \left\| I - K^{-1}\tilde{K} \right\|_2 + \left\| \left(I - K^{-1}\tilde{K} \right)^2 \right\|_2 + \cdots\cdots \\
&\leq 1 + \frac{\varepsilon}{\varepsilon_0} + \left(\frac{\varepsilon}{\varepsilon_0} \right)^2 + \cdots\cdots \\
&= \frac{\varepsilon_0}{\varepsilon_0 - \varepsilon}.
\end{aligned}$$

Then, we can obtain

$$\left\| \tilde{K}^{-1} \right\|_2 = \left\| \tilde{K}^{-1} K K^{-1} \right\|_2 \leq \left\| \tilde{K}^{-1} K \right\|_2 \left\| K^{-1} \right\|_2 \leq \frac{\varepsilon_0}{\varepsilon_0 - \varepsilon} \times \frac{1}{\varepsilon_0} = \frac{1}{\varepsilon_0 - \varepsilon}.$$

This in turn yields

$$\left\| I - \tilde{K}^{-1} K \right\|_2 = \left\| \tilde{K}^{-1}(\tilde{K} - K) \right\|_2 \leq \left\| \tilde{K}^{-1} \right\|_2 \left\| \tilde{K} - K \right\|_2 \leq \frac{\varepsilon}{\varepsilon_0 - \varepsilon}. \qquad \square$$

According to Theorem 3.6 we have the following proposition that gives the condition number of the preconditioned matrix.

PROPOSITION 3.7. *Let a nonsingular matrix K be of the form* (2.2) *and let the condition ε in Definition 3.1 hold by Corollary 3.5. If $\tilde{S}_i$ $(i = 1, 2, \ldots, N)$ are*

nonsingular and $\varepsilon < \frac{1}{2}\varepsilon_0$*, then we have the MSSS preconditioner* $\tilde{K}$ *is of the form* (2.4) *and*

$$\kappa_2(\tilde{K}^{-1}K) \leq \frac{\varepsilon_0}{\varepsilon_0 - 2\varepsilon}.$$

Here ε_0 *is the smallest singular value of* K*.*

Proof. According to Theorem 3.6, we have

$$\left\|I - \tilde{K}^{-1}K\right\|_2 \leq \frac{\varepsilon}{\varepsilon_0 - \varepsilon},$$

associated with $\varepsilon < \frac{1}{2}\varepsilon_0$, we get $\dfrac{\varepsilon}{\varepsilon_0 - \varepsilon} < 1$. Then the Neumann series

$$I + \left(I - \tilde{K}^{-1}K\right) + \left(I - \tilde{K}^{-1}K\right)^2 + \cdots\cdots$$

converge to $\left(I - (I - \tilde{K}^{-1}K)\right)^{-1} = K^{-1}\tilde{K}$. This yields

$$\begin{aligned}
\left\|K^{-1}\tilde{K}\right\|_2 &= \left\|I + \left(I - \tilde{K}^{-1}K\right) + \left(I - \tilde{K}^{-1}K\right)^2 + \cdots\cdots\right\|_2 \\
&\leq 1 + \left\|I - \tilde{K}^{-1}K\right\|_2 + \left\|\left(I - \tilde{K}^{-1}K\right)^2\right\|_2 + \cdots\cdots \\
&\leq 1 + \frac{\varepsilon}{\varepsilon_0 - \varepsilon} + \left(\frac{\varepsilon}{\varepsilon_0 - \varepsilon}\right)^2 + \cdots\cdots \\
&= \frac{\varepsilon_0 - \varepsilon}{\varepsilon_0 - 2\varepsilon}.
\end{aligned}$$

According to Theorem 3.6, we have

$$\left\|\tilde{K}^{-1}K\right\|_2 \leq \frac{\varepsilon_0}{\varepsilon_0 - \varepsilon},$$

then we obtain

$$\kappa_2(\tilde{K}^{-1}K) = \left\|\tilde{K}^{-1}K\right\|_2 \left\|K^{-1}\tilde{K}\right\|_2 \leq \frac{\varepsilon_0}{\varepsilon_0 - 2\varepsilon} \qquad \square$$

According to Theorem 3.6, we can also give an analytic bound on the spectrum of the preconditioned matrix.

PROPOSITION 3.8. *Let a nonsingular matrix* K *be of the form* (2.2) *and let the condition* ε *in Definition 3.1 hold by following Corollary 3.5. If* $\tilde{S}_i$ $(i = 1, 2, \ldots, N)$ *are nonsingular, then we have the MSSS preconditioner* $\tilde{K}$ *is of the form* (2.4)*. Denote the eigenvalues of the preconditioned matrix by* $\lambda(\tilde{K}^{-1}K)$*. If* $\varepsilon < \varepsilon_0$*, we have*

$$\left|\lambda(\tilde{K}^{-1}K) - 1\right| \leq \frac{\varepsilon}{\varepsilon_0 - \varepsilon}.$$

Here ε_0 *is the smallest singular value of* K*.*

Proof. According to Theorem 3.6, we have

$$\left\|I - \tilde{K}^{-1}K\right\|_2 \leq \frac{\varepsilon}{\varepsilon_0 - \varepsilon}.$$

Therefore, we can obtain

$$\left|\lambda(I-\tilde{K}^{-1}K)\right| \leq \frac{\varepsilon}{\varepsilon_0-\varepsilon},$$

owing to $\left|\lambda(I-\tilde{K}^{-1}K)\right| \leq \left\|I-\tilde{K}^{-1}K\right\|_2$. Since $\lambda(I-\tilde{K}^{-1}K) = 1-\lambda(\tilde{K}^{-1}K)$, we get

$$\left|\lambda(\tilde{K}^{-1}K)-1\right| \leq \frac{\varepsilon}{\varepsilon_0-\varepsilon}. \qquad \square$$

Remark 3.2. Proposition 3.8 states that the spectrum of the preconditioned system is contained in a circle centered at $(1,0)$ with a maximum radius $\frac{\varepsilon}{\varepsilon_0-\varepsilon}$. Therefore, the smaller ε is, the closer the eigenvalues of the preconditioned system are to $(1,0)$. This in turn gives better convergence for a wide class of Krylov solvers to solve the preconditioned system by applying the MSSS preconditioner.

According to Theorem 3.6, Proposition 3.7, and Proposition 3.8, we conclude that the smaller the ε is, the better-conditioned the preconditioned matrix is. For the extreme case $\varepsilon = 0$ when there is no approximation of the Schur complement, this factorization is exact. This is in turn verified by Theorem 3.6, Proposition 3.7, and Proposition 3.8. In Section 5, we will show that ε can be made arbitrarily small by setting a parameter in the MSSS preconditioner.

4. Breakdown Free Condition. In the previous section, we have analyzed the conditioning and spectrum of the preconditioned matrix. In this section, we discuss how to compute the MSSS preconditioner without breakdown, i.e., how to set the bound ε to compute the nonsingular $\tilde{S}_i$ $(i = 1,\ 2,\ \ldots,\ N)$. To start with, we give the following lemmas that are necessary for the analysis.

Lemma 4.1 ([40]). *Let A be an $m\times n$ matrix with, say, $m \geq n$. Sort its singular values in a non-increasing order by*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n.$$

Let $\tilde{A} = A+E$ be a perturbation of A, and sort its singular values in a non-increasing order by

$$\tilde{\sigma}_1 \geq \tilde{\sigma}_2 \geq \cdots \geq \tilde{\sigma}_n.$$

Then, we have

$$|\tilde{\sigma}_i - \sigma_i| \leq \|E\|_2, \quad i = 1,\ 2,\ \cdots,\ n.$$

By applying this lemma, we have the following lemma that gives a sufficient condition for a nonsingular perturbation, i.e., for a full rank matrix A, its perturbed analog $\tilde{A}$ is still of full rank.

Lemma 4.2. *Let A be an $m\times n$ full rank matrix with, say, $m \geq n$, and $\tilde{A} = A+E$ be a perturbation of A. If*

$$\|E\|_2 < \sigma_n,$$

where σ_n is the smallest singular value of A, then the perturbed matrix $\tilde{A}$ is still of full rank.

Proof. Denote the smallest singular value of $\tilde{A}$ by $\tilde{\sigma}_n$, then according to Lemma 4.1, we have

$$|\sigma_n - \tilde{\sigma}_n| \leq \|E\|_2 .$$

Since $\|E\|_2 < \sigma_n$, this yields,

$$|\sigma_n - \tilde{\sigma}_n| < \sigma_n.$$

We can obtain

$$0 < \tilde{\sigma}_n < 2\sigma_n,$$

which states that $\tilde{A}$ is still of full rank. ∎

With these lemmas, we can give a sufficient condition for ε that guarantees nonsingular approximate Schur complements $\tilde{S}_i (i = 1,\ 2,\ \ldots,\ N)$, which satisfies condition ε in Definition 3.1.

THEOREM 4.3. *If ε satisfies the following inequality,*

$$\varepsilon < \sigma_0$$

where

$$\sigma_0 \triangleq \min_{p=1}^{N} \Big\{ \min \Big(\sigma(K_p) \Big) \Big\},$$

and K_p is the leading principle sub-matrix of K of size $pN \times pN$ for $p = 1,\ 2,\ \cdots,\ N$. Here $\min(\sigma(K_p))$ denotes the smallest singular value of K_p. Then the condition ε can be satisfied and all the approximate Schur complements $\tilde{S}_i$ $(i = 1,\ 2,\ \ldots,\ N)$ are nonsingular.

Proof. At each step j $(j \geq 2)$ of the approximate factorization, we use $\tilde{S}_j$ to approximate $K_{j,j} - K_{j,j-1}\tilde{S}_{j-1}^{-1}K_{j-1,j}$ by performing a model order reduction according to Corollary 3.5 to satisfy the condition ε. This introduces a small perturbation that is given by

$$E_j \triangleq \tilde{S}_j - \Big(K_{j,j} - K_{j,j-1}\tilde{S}_{j-1}^{-1}K_{j-1,j} \Big),$$

and $\|E_j\|_2 \leq \varepsilon$ $(j \geq 2)$. Since $S_1 = K_{1,1}$ is an SSS matrix with small off-diagonal rank, no model order reduction is performed, we have $E_1 = 0$.

Denote the $(j-1)$-th leading principle sub-matrix of K by $\bar{K}_{j-1}$, then we have

$$\bar{K}_j = \begin{bmatrix} \bar{K}_{j-1} & \bar{K}_{j-1,j} \\ \bar{K}_{j,j-1} & K_{j,j} \end{bmatrix}, \quad \tilde{\bar{K}}_j = \begin{bmatrix} \tilde{\bar{K}}_{j-1} & \bar{K}_{j-1,j} \\ \bar{K}_{j,j-1} & K_{j,j} + E_j \end{bmatrix}, \tag{4.1}$$

where $\bar{K}_j$ is the j-th principle leading sub-matrix of K and $\tilde{\bar{K}}_j$ is its approximation. $\tilde{\bar{K}}_j$ is also the j-th leading principle sub-matrix of $\tilde{K}$, and $\bar{K}_{j,j-1} = \begin{bmatrix} \mathbf{0} & K_{j,j-1} \end{bmatrix}$, $\bar{K}_{j-1,j} = \begin{bmatrix} \mathbf{0} \\ K_{j-1,j} \end{bmatrix}$. Moreover, we have

$$\tilde{\bar{K}}_j - \bar{K}_j = \begin{bmatrix} E_1 & & & \\ & E_2 & & \\ & & \ddots & \\ & & & E_j \end{bmatrix},$$

where $\|E_i\|_2 \leq \varepsilon$ $(1 \leq i \leq j)$. Then we can obtain,

$$\begin{aligned}\left\|\left(\tilde{\bar{K}}_j - \bar{K}_j\right)x\right\|_2^2 &= \|E_1x_1\|_2^2 + \|E_2x_2\|_2^2 + \cdots + \|E_jx_j\|_2^2 \\ &\leq \|E_1\|_2^2\|x_1\|_2^2 + \|E_2\|_2^2\|x_2\|_2^2 + \cdots + \|E_j\|_2^2\|x_j\|_2^2 \\ &\leq \varepsilon^2\sum_{k=1}^{j}\|x_k\|_2^2 = \varepsilon^2\|x\|_2^2.\end{aligned}$$

This in turn yields

$$\left\|\tilde{\bar{K}}_j - \bar{K}_j\right\|_2 \leq \varepsilon,\ j = 1, 2, \cdots, N.$$

According to Lemma 4.2, if $\epsilon < \min\limits_{j=1}^{N}\left\{\min\left(\sigma(\bar{K}_j)\right)\right\}$, then $\tilde{\bar{K}}_j$ $(j = 1, 2, \cdots, N)$ is nonsingular.

Since $\tilde{\bar{K}}_j$ and $\tilde{\bar{K}}_{j-1}$ are nonsingular, according to (4.1), the Schur complement of $\tilde{\bar{K}}_{j-1}$ in $\tilde{\bar{K}}_j$ is also nonsingular and is given by

$$\tilde{\bar{S}}_j = K_{j,j} + E_j - \bar{K}_{j,j-1}\tilde{\bar{K}}_{j-1}^{-1}\bar{K}_{j-1,j},$$

which is exactly $\tilde{S}_j$ for $j = 2, \cdots, N$, while $\tilde{S}_1 = K_{1,1}$ is also nonsingular. □

Theorem 4.3 gives a sufficient condition for ε to compute nonsingular approximate Schur complements for a general class of linear systems. For the symmetric positive definite case, this sufficient condition can be simplified by the following lemma.

LEMMA 4.4. *Let K be an $m \times m$ symmetric positive definite matrix of the form* (2.2) *and denote its smallest eigenvalue by $\lambda_{\min}(K)$. If $\varepsilon < \lambda_{\min}(K)$, then all the approximated Schur complements $\tilde{S}_i$ $(i = 1,\ 2,\ \ldots,\ N)$ are nonsingular.*

Before giving the proof of Lemma 4.4, we first introduce the following lemma that is necessary for the proof.

LEMMA 4.5 (Corollary 8.4.6 in [4]). *Let A be an $m \times m$ Hermitian matrix, and A_0 be a $k \times k$ principle sub-matrix of A with $k < m$. Then,*

$$\lambda_{\min}(A) \leq \lambda_{\min}(A_0) \leq \lambda_{\max}(A_0) \leq \lambda_{\max}(A),$$

and

$$\lambda_{\min}(A_0) \leq \lambda_k(A).$$

With Lemma 4.5, we give the proof of Lemma 4.4 as follows.

Proof. According to Theorem 4.3, if $\varepsilon < \min\limits_{p=1}^{N}\left\{\min\left(\sigma(K_p)\right)\right\}$, the approximate Schur complements $\tilde{S}_i$ $(i = 1,\ 2,\ \ldots,\ N)$ are nonsingular. For the symmetric positive definite matrix K, its eigenvalues and singular values are identical. Then the condition for ε is given by

$$\varepsilon < \min_{p=1}^{N}\left\{\min\left(\lambda(K_p)\right)\right\}.$$

According to Lemma 4.5, we have

$$\min\left(\lambda(K_p)\right) \geq \lambda_{\min}(K),$$

this in turn gives the condition for ε, that is

$$\varepsilon < \lambda_{\min}(K). \qquad \square$$

Theorem 4.3 gives a sufficient condition for ε to obtain nonsingular approximate Schur complements. Next, we use a simple example to illustrate this condition.

EXAMPLE 4.1. *Consider the 2D stationary SchröĬĹdinger equation*

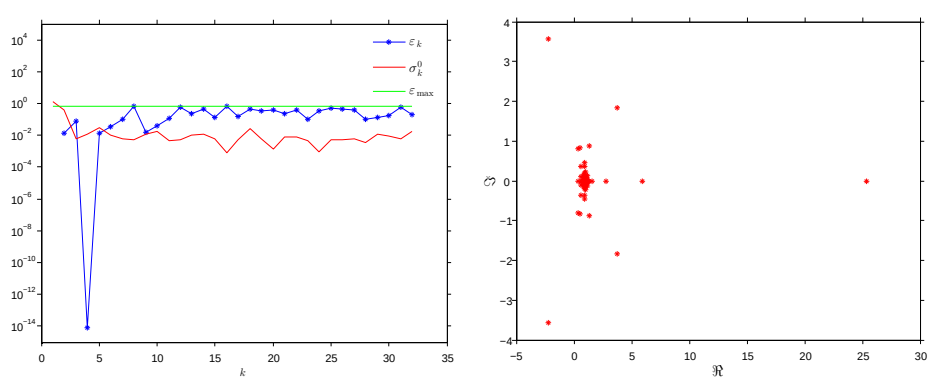
$$\nabla^2\Psi(x,y) + k^2\Psi(x,y) = 0,$$

with homogeneous Dirichlet boundary condition on a unite square domain $\Omega = [0,\ 1]\times[0,\ 1]$. *Using* 5*-point stencil finite difference discretization on a uniform grid with gird size* $h = 2^{-5}$ *gives a linear system that has a* 2*-level SSS structure, here* $k^2h^2 = {}^{\pi^2}/_{16}$. *Factorize the linear system by using MSSS matrix computations that satisfies condition* ε *in Definition 3.1 gives an MSSS preconditioner* $\tilde{K}$.

For different settings of ε, the smallest singular value σ_k^0 of the leading principle sub-matrix K_k of size $kN \times kN$, the approximation error ε_k at each step to compute the approximate Schur complement, the bound of ε_k which is denoted by $\varepsilon_{\max}$, and the preconditioned spectrum are plotted in Figure 1-3.

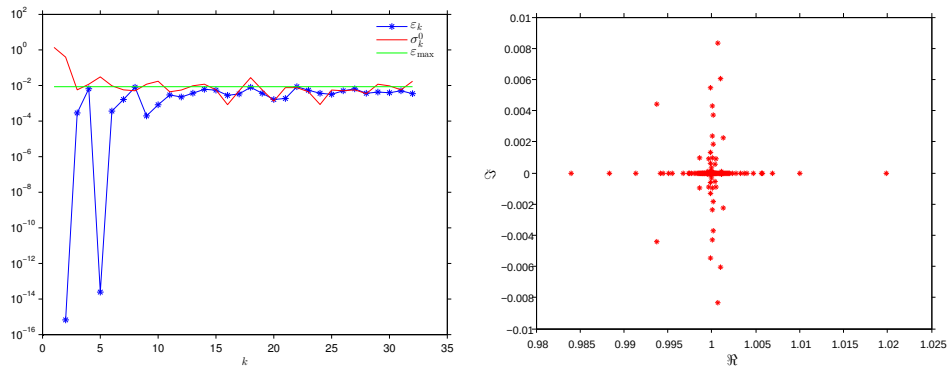
We start decreasing ε from 0.5 to 10^{-3}, this corresponds to $\varepsilon > \sigma_0$ for relatively big ε. For the case $\varepsilon > \sigma_0$, Theorem 4.3 does not hold, which means that we may fail to get nonsingular approximate Schur complements $\tilde{S}_i$. However, we succeed in computing the nonsingular Schur complements $\tilde{S}_i$. In fact, the possibility of perturbing a matrix from nonsingularity to singularity is quite small. Although we get nonsingular approximate Schur complements for $\varepsilon > \sigma_0$, our analysis is not suited to analyzing the preconditioned spectrum for such case. This is illustrated by the spectrum of the preconditioned system in Figure 1(b). The preconditioned spectrum corresponds to $\varepsilon = \mathcal{O}(0.5)$ is not well clustered and a portion of the eigenvalues is far away from $(1, 0)$.

For the cases that ε is slightly bigger than σ_0, the preconditioned spectrum is already contained in a quite small circle, cf. Figure 2(b). When ε is of the same order as σ_0, the circle is even smaller, cf. Figure 2(b) and Figure 3(b). Continue decreasing ε, the radius of the circle can be made arbitrarily small. At a certain moment, the MSSS factorization can be used as a direct solver for small enough ε.



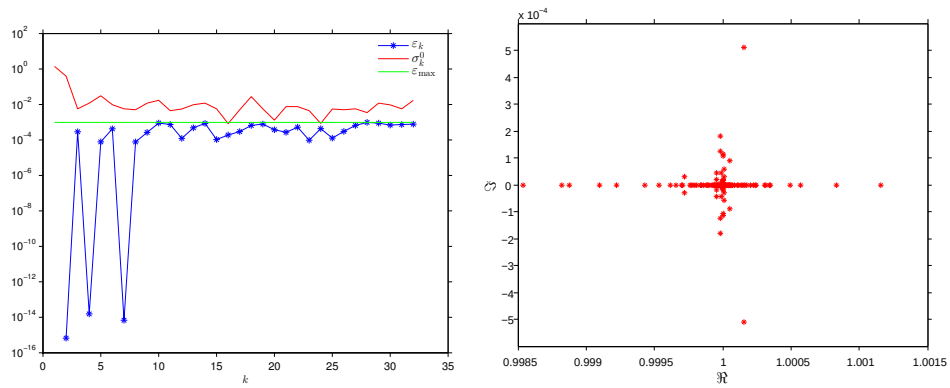
(a) Smallest singular values and ε_k (b) Preconditioned spectrum

FIG. 1. *Condition* ε *and preconditioned spectrum for* $\varepsilon = \mathcal{O}(0.5)$



(a) Smallest singular values and ε_k (b) Preconditioned spectrum

FIG. 2. *Condition ε and preconditioned spectrum for $\varepsilon = \mathcal{O}(10^{-2})$*



(a) Smallest singular values and ε_k (b) Preconditioned spectrum

FIG. 3. *Condition ε and preconditioned spectrum for $\varepsilon = \mathcal{O}(10^{-3})$*

REMARK 4.1. We observed that for the case $\varepsilon < \sigma_0$ and ε is of the same order as σ_0, if ε decreases by a factor 10, the radius of the circle that contains the preconditioned spectrum is also reduced by a factor of around 10. This verifies the statement of the radius of the circle in Proposition 3.8.

The results for different ε given by Figure 1-3 verify our analysis of the convergence property of the MSSS preconditioner and the spectrum of the preconditioned system in Section 3. Both theoretical analysis and numerical experiments state that a small ε is preferred. Normally, decreasing ε will increase the computational complexity to a small extent. It is therefore favorable to choose a moderate ε to compute an MSSS factorization and use it as a preconditioner. This gives linear computational complexity and satisfactory convergence for a wide class of Krylov solvers. Details will be discussed in Section 6.

Both theoretical analysis and numerical experiments indicate that a small ε is preferred. In the next section, we will discuss how to perform the model order reduction to make ε up to a prescribed accuracy.

5. Approximation Error of Model Order Reduction. We assume by Corollary 3.5 that the condition ε is satisfied via a model order reduction operation and the error of the model order reduction should be bounded by ε. To start, we use the model order reduction algorithm which is called Hankel blocks approximation that is studied in [8]. In the following part, we will show how to do this model order reduction

to make the approximation error up to a prescribed accuracy ε. In this section, we use the algorithm style notation, i.e., by letting $a = b$, we assign the variable a with the value of b. To begin this analysis, we recall some concepts that are necessary.

Definition 5.1 ([8]). *Hankel blocks denote the off-diagonal blocks that extend from the diagonal to the northeast corner (for the upper case) or to the southwest corner (for the lower case).*

Take a 4×4 block SSS matrix A for example, the Hankel blocks for the strictly lower-triangular part are shown in Figure 4 by $\mathcal{H}_2$, $\mathcal{H}_3$ and $\mathcal{H}_4$.

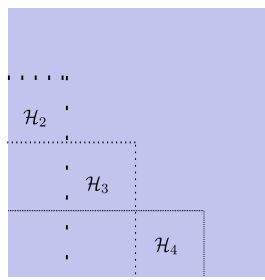


Fig. 4. *Hankel blocks of a 4×4 block SSS matrix*

For the Hankel blocks $\mathcal{H}_k$ of the SSS matrix A, it has the following low-rank factorization,

$$\mathcal{H}_k = \mathcal{O}_k \mathcal{C}_k,$$

where the low-rank factors $\mathcal{O}_k$ and $\mathcal{C}_k$ have a backward and forward recursion, respectively. They are given by

$$\begin{cases} \mathcal{O}_k &= P_N, \text{ if } \quad k = N, \\ \mathcal{O}_k &= \begin{bmatrix} P_k \\ \mathcal{O}_{k+1} R_k \end{bmatrix}, \text{ if } \quad 2 \le k < N, \end{cases}$$

and

$$\begin{cases} \mathcal{C}_k &= Q_1, \text{ if } \quad k = 2, \\ \mathcal{C}_k &= \begin{bmatrix} R_{k-1}\mathcal{C}_{k-1} & Q_{k-1} \end{bmatrix}, \text{ if } \quad 2 < k \le N. \end{cases}$$

The rank r_k of the Hankel block $\mathcal{H}_k$ has the following equality

$$\operatorname{rank}(\mathcal{H}_k) = \operatorname{rank}(\mathcal{O}_k) = \operatorname{rank}(\mathcal{C}_k) = r_k.$$

The low-rank factors $\mathcal{O}_k$ and $\mathcal{C}_k$ are the observability factor and controllability factor of a linear time-varying (LTV) system that corresponds to the SSS matrix A. Moreover, the Hankel block $\mathcal{H}_k$ corresponds to the discrete Hankel map of a LTV system. SSS matrices and their relations with LTV system are studied in [12].

The basic idea for the model order reduction of SSS matrices is to reduce the rank of the Hankel blocks $\mathcal{H}_k$ from r_k to $\tilde{r}_k$ with $\tilde{r}_k < r_k$, where $\tilde{r}_k$ is the rank of the approximated Hankel map $\tilde{\mathcal{H}}_k$ and

$$\operatorname{rank}(\tilde{\mathcal{H}}_k) = \operatorname{rank}(\tilde{\mathcal{O}}_k) = \operatorname{rank}(\tilde{\mathcal{C}}_k) = \tilde{r}_k.$$

To start the model order reduction, first we need to transform $\mathcal{C}_i$ to the form that has orthonormal rows, which is called the right-proper form in [8]. This is obtained by performing a singular value decomposition (SVD) on $\mathcal{C}_i$. For $i = 2$,

$$\mathcal{C}_2 = U_1 \Sigma_1 V_1^T,$$

and let $\mathcal{C}_2 = Q_1 = V_1^T$. To keep the Hankel map (block) $\mathcal{H}_2$ unchanged, we let $\mathcal{O}_2 = \mathcal{O}_2 U_1 \Sigma_1$. This gives

$$P_2 = P_2 U_1 \Sigma_1, \ R_2 = R_2 U_1 \Sigma_1.$$

From step i to $i+1$, we have

$$\mathcal{C}_{i+1} = \begin{bmatrix} R_i \mathcal{C}_i & Q_i \end{bmatrix} = \begin{bmatrix} R_i & Q_i \end{bmatrix} \begin{bmatrix} \mathcal{C}_i & \\ & I \end{bmatrix}.$$

Since $\mathcal{C}_i$ has orthonormal rows, $\begin{bmatrix} \mathcal{C}_i & \\ & I \end{bmatrix}$ also has orthonormal rows. To complete this procedure for $\mathcal{C}_{i+1}$, perform the following SVD

$$\begin{bmatrix} R_i & Q_i \end{bmatrix} = U_i \Sigma_i V_i^T,$$

and let $\begin{bmatrix} R_i & Q_i \end{bmatrix} = V_i^T$. To keep the Hankel map at step $i+1$ unchanged, we let $\mathcal{O}_{i+1} = \mathcal{O}_{i+1} U_i \Sigma_i$. After finishing the above procedure, we can make all the factors $\mathcal{C}_i$ have orthonormal rows.

The next step of the model order reduction is to transform the low-rank factors $\mathcal{O}_i$ to the form with orthonormal columns, which is called the left-proper form. Then we reduce the rank of the Hankel map (blocks). Since the recursion for the factor $\mathcal{O}_i$ is performed backward, we start from $i = N$.

First we approximate $\mathcal{O}_N$ by $\tilde{\mathcal{O}}_N$, this gives the factor $\tilde{\mathcal{O}}_{N-1}^1$ for the next step. Here $\tilde{\mathcal{O}}_{N-1}^1$ denotes the approximated factor $\mathcal{O}_{N-1}$ because of the propagation of the approximation error of $\mathcal{O}_N$. Then we compute a low-rank approximation of $\mathcal{O}_{N-1}^1$, which gives $\tilde{\mathcal{O}}_{N-1}^2$. We continue this procedure till step $i = 2$. At step i, we use $\tilde{\mathcal{O}}_i^2$ to approximate $\tilde{\mathcal{O}}_i^1$, this introduces an approximation error that is bounded by τ. We use Figure 5 to depict this backward recursion of approximation. For the details of the Hankel blocks approximation algorithm, cf. [8, 31].

$$\begin{array}{ccccccccc}
\mathcal{O}_N & \longrightarrow & \mathcal{O}_{N-1} & \longrightarrow & \mathcal{O}_{N-2} & \cdots & \longrightarrow & \mathcal{O}_2 \\
\downarrow \tau & & & & & & & \\
\tilde{\mathcal{O}}_N & \rightarrow & \tilde{\mathcal{O}}_{N-1}^1 & \xrightarrow{\tau} & \tilde{\mathcal{O}}_{N-1}^2 & \rightarrow & \tilde{\mathcal{O}}_{N-2}^1 \xrightarrow{\tau} \tilde{\mathcal{O}}_{N-2}^2 & \cdots \tilde{\mathcal{O}}_2^2
\end{array}$$

FIG. 5. *Low-rank approximation of $\mathcal{O}_k$*

Here $\tilde{\mathcal{O}}_k^1$ represents the approximation of $\mathcal{O}_k$ by considering the propagation of the error introduced in the previous steps. And $\tilde{\mathcal{O}}_k^1$ is further approximated by $\tilde{\mathcal{O}}_k^2$ by performing a low-rank approximation. Then we have the following lemma that underlies the error between the original Hankel map and the approximate Hankel map.

LEMMA 5.2. *Let A be a block $N \times N$ SSS matrix and its Hankel blocks $\mathcal{H}_i$ be approximated by $\tilde{\mathcal{H}}_i$ using the Hankel blocks approximation described by the above procedure, then*

$$\left\| \mathcal{H}_i - \tilde{\mathcal{H}}_i \right\|_2 \leq (N - i + 1)\tau, \quad 2 \leq i \leq N. \tag{5.1}$$

where τ is the upper bound for the discarded singular values that is applied in the singular value decomposition of the Hankel factors. For the approximated Hankel factors $\mathcal{O}_i$ that are illustrated in Figure 5, we have

$$\left\|\mathcal{O}_i - \tilde{\mathcal{O}}_i^1\right\|_2 \leq (N-i)\tau, \quad 2 \leq i \leq N-1. \tag{5.2}$$

and

$$\left\|\tilde{\mathcal{O}}_i^2 \tilde{Q}_{i-1} - \mathcal{O}_i Q_{i-1}\right\|_2 \leq (N-i+1)\tau, \quad 2 \leq i \leq N. \tag{5.3}$$

Here we use the "˜" notation to denote a factor or matrix after approximation.

Proof. For the proof of this lemma, cf. Appendix A. □

REMARK 5.1. To perform the Hankel blocks approximation to reduce the off-diagonal rank of an SSS matrix, the reduction of the strictly lower-triangular part is clearly covered by the analysis above. To reduce the off-diagonal rank of the strictly upper-triangular part, we can first transpose it to the strictly lower-triangular form. Then perform the Hankel blocks approximation to the strictly lower-triangular part and transpose back to the strictly upper-triangular form. This gives strictly upper-triangular part with reduced off-diagonal rank.

In Section 3, we gave an analytic bound of the radius of the circle that contains the preconditioned spectrum. This analytic bound is closely related to the approximation error ε by the model order reduction for SSS matrices, cf. Corollary 3.5 and Proposition 3.8. This model order reduction error ε can be made arbitrarily small by setting a parameter in the MSSS preconditioner. Now, we have all the ingredients to help to compute a controllable ε. Next, we will give the main theorem of this section to show how to compute the controllable ε.

THEOREM 5.3. *Let the $N \times N$ block SSS matrix A be approximated by $\tilde{A}$ with lower off-diagonal rank using the Hankel blocks approximation, then*

$$\left\|A - \tilde{A}\right\|_2 \leq 2\sqrt{N}(N-1)\tau,$$

where τ is the upper bound of the discarded singular values for the singular value decomposition that is performed in approximating the Hankel blocks.

Proof. To prove this theorem, we use Figure 6 to illustrate the column structure of the off-diagonal blocks for an SSS matrix. Since the strictly upper-triangular part and the strictly lower-triangular part have similar structure, here we just take the strictly lower-triangular part for example.

It is not difficult to verify that the i-th off-diagonal column of the strictly lower triangular part of an SSS matrix, denoted by C_i, can be represented by

$$C_i = \mathcal{O}_{i+1} Q_i, \quad (i = 1, 2, \cdots, N-1).$$

After performing the Hankel blocks approximation, C_i is approximated by

$$\tilde{C}_i = \tilde{\mathcal{O}}_{i+1}^2 \tilde{Q}_i, \quad (i = 1, 2, \cdots, N-1).$$

(a) Hankel columns (b) Approximate Hankel columns

FIG. 6. *Lower-triangular Hankel columns before and after approximation*

Denote $\Delta C_i = C_i - \tilde{C}_i$, then we have

$$\|\Delta C_i\|_2 = \left\|\mathcal{O}_{i+1}Q_i - \tilde{\mathcal{O}}_{i+1}^2\tilde{Q}_i\right\|_2 \leq (N-i)\tau. \quad \text{(Lemma 5.2)}$$

We can write the SSS matrix A by using the following form

$$A = L + D + U,$$

where L is the strictly lower-triangular part, D is the block diagonal part, and U is the strictly upper-triangular part of A, respectively. Performing the Hankel blocks approximation on the strictly lower-triangular part and the strictly upper-triangular part, we obtain

$$\tilde{A} = \tilde{L} + D + \tilde{U},$$

where $\tilde{L}$ and $\tilde{U}$ are approximated independently. Moreover, we have

$$\tilde{L} - L = \begin{bmatrix}\Delta C_1 & \Delta C_2 & \cdots & \Delta C_{N-1} & \mathbf{0}\end{bmatrix}.$$

And

$$\begin{aligned}\left\|\left(\tilde{L} - L\right)x\right\|_2 &= \left\|\sum_{i=1}^{N-1}\Delta C_i x_i\right\|_2 \leq \sum_{i=1}^{N-1}\|\Delta C_i\|_2\|x_i\|_2 \leq \max_{i=1}^{N-1}\|\Delta C_i\|_2\sum_{i=1}^{N-1}\|x_i\|_2 \\ &\leq (N-1)\tau\sum_{i=1}^{N-1}\|x_i\|_2 \leq (N-1)\tau\sum_{i=1}^{N}\|x_i\|_2 \\ &\leq (N-1)\tau\sqrt{N\sum_{i=1}^{N}\|x_i\|_2^2} = \sqrt{N}(N-1)\tau\|x\|_2.\end{aligned}$$

This yields

$$\left\|L - \tilde{L}\right\|_2 \triangleq \max_{x\neq\mathbf{0}}\frac{\left\|(L-\tilde{L})x\right\|_2}{\|x\|_2} \leq \sqrt{N}(N-1)\tau.$$

Here τ is the upper bound for the discarded singular values for the singular value decomposition that is performed in the Hankel blocks approximation. Similarly, we have $\left\|U - \tilde{U}\right\|_2 \leq \sqrt{N}(N-1)\tau$, this gives

$$\begin{aligned}\left\|A - \tilde{A}\right\|_2 &= \left\|(L-\tilde{L}) + (U-\tilde{U})\right\|_2 \\ &\leq \left\|L - \tilde{L}\right\|_2 + \left\|U - \tilde{U}\right\|_2 \qquad \square \\ &\leq 2\sqrt{N}(N-1)\tau.\end{aligned}$$

Remark 5.2. Theorem 5.3 gives an analytical bound of the error introduced by the model order reduction. This error is only related to the maximum discarded singular value τ for the singular value decomposition that is performed in the Hankel blocks approximation. This states that this approximation error can be made arbitrarily small by setting τ small enough. This in turn gives a relatively bigger off-diagonal rank compared with moderate τ. A trade-off has to be made between the computational complexity and accuracy.

Remark 5.3. The model order reduction can be also performed by setting a fixed reduced off-diagonal rank. This is convenient in practice and makes the computational complexity and memory consumption easily predictable. The disadvantage, however, is that the approximation error is unpredictable. In contrast, by setting a fixed τ for the model order reduction, we can easily control the error bound and get an adaptive reduced off-diagonal rank. However, the disadvantage is that the computational complexity is difficult to estimate. In practice, we observe that for many applications, a properly chosen τ also gives small enough off-diagonal rank, which in turn gives predictable computational complexity. This will be highlighted in the numerical experiments part.

Remark 5.4. In many applications, the error introduced by the model order reduction using the Hankel blocks approximation is of $\mathcal{O}(\tau)$, which is quite small compared with the bound given by Theorem 5.3. Only in some extreme cases, the bound given by Theorem 5.3 is sharp. However, it is quite difficult to prove for which case the error bound given by Theorem 5.3 is sharp. Normally, a small τ still results a small reduced off-diagonal rank, which yields linear computational complexity. This will be illustrated by numerical experiments in the next section.

In practice, it would be desirable to estimate the semiseparable order of the Schur complement that corresponds to a given τ. Normally, this is quite challenging since the off-diagonal rank depends not only on the differential operator of the PDE, but also on the coefficients of the PDE. Only some preliminary results can be found in the literature. These results are summarized by Lemma 5.4.

Lemma 5.4 ([9]). *Let the symmetric positive definite block tridiagonal system K arise from the discretization of PDEs with Laplacian operator, constant coefficients, and Dirichlet boundary condition everywhere on the boundary. Then the Schur complement S_i has a monotonically convergence rate and the limit of S_i, i.e., S_∞ is also symmetric positive definite. The τ-rank of the Hankel blocks of S_∞ are bounded by*

$$r\left(1+8\ln^4\left(\frac{3\,\|D\|}{\tau}\right)\right),$$

where r is the maximal Hankel block rank of $K_{i,i}$ and $K_{i,i-1}$, D is the diagonal block of K. Here, the τ-rank of a matrix is defined by the number of singular values that are bigger than or equal to τ.

Lemma 5.4 gives the upper bound of the limit of the Schur complement for the infinite dimensional symmetric positive definite systems. For finite dimensional symmetric positive definite systems with constant coefficients, similar results hold. For detailed discussion, cf. [9]. Note that this bound is not sharp because the term $\ln^4\left(\frac{3\,\|D\|}{\tau}\right)$ can be much bigger than the size of K.

Recall Lemma 4.4 states that for the symmetric positive definite system K, τ can be often chosen as $\tau < \lambda_{\min}(K)$. In Example 4.1, it is shown that usually we can choose $\tau = \mathcal{O}(\lambda_{\min}(K))$. If $\|D\| = \mathcal{O}(\|K\|)$, then we get the bound of the rank of the

Hankel blocks is of $\mathcal{O}\left(r \ln^4 \kappa_2(K)\right)$. Even this bound is not sharp, it states that for ill-conditioned system, a bigger semiseparable order is needed to get a considerably good approximation, which will be shown in Section 6.

For the symmetric positive definite systems from the discretization of PDEs with variable coefficients and the indefinite systems, the analytic bound of the rank of the Hankel blocks of the Schur complement is quite difficult to analyze. Relevant work on analyzing the off-diagonal rank of the Schur complement of the symmetric positive definite type by using hierarchical matrix computations is done in [2].

Remark 5.5. The τ rank of the off-diagonal blocks of the Schur complement for symmetric positive definite systems studied in Lemma 5.4 is not sharp. In many applications, it can be made quite small and even bounded by a small number for a wide class of linear systems. This will be illustrated by numerical experiments in the next section.

6. Numerical Experiments. In this section, we use numerical experiments to investigate our analysis in the previous sections. We use three types of experiments, which include unsymmetric systems, symmetric indefinite systems from discretization of scalar PDEs and saddle-point systems, to demonstrate our results. For all the numerical experiments performed in this section, the induced dimension reduction (`IDR(s)`)[44] is used as a Krylov solver. The `IDR(s)` solver is terminated when the 2-norm of the residual is reduced by a factor of 10^{-6}. The numerical experiments are implemented in MATLAB 2011b on a desktop of Intel Core i5 CPU of 3.10 GHz and 16 Gb memory with the Debian GNU/Linux 8.0 system.

6.1. Unsymmetric System. In this subsection, we use the convection-diffusion equation as a numerical example to demonstrate our analysis for the unsymmetric case. The convection-diffusion problem is described by Example 6.1, which is given as the example 3.1.4 in [15]. The details of the discretization of the convection-diffusion equation can be also found in [15]. We generate the linear system in this example using the Incompressible Flow and Iterative Solver Software [38] (IFISS) *. To investigate the performance of the MSSS preconditioning technique, we consider the case for a moderate ν and a small ν.

Example 6.1 ([15]). *Zero source term, recirculating wind, characteristic boundary layers.*

$$\begin{aligned} -\nu\nabla^2 u + \overrightarrow{\omega}\cdot\nabla u &= f \quad \text{in} \quad \Omega \\ u &= u_D \quad \text{on} \quad \partial\Omega \end{aligned} \tag{6.1}$$

where $\Omega = \{(x,y)| -1 \le x \le 1, -1 \le y \le 1\}$, $\overrightarrow{\omega} = \left(2y(1-x^2),\ -2x(1-y^2)\right)$, $f = 0$. *Homogeneous Dirichlet boundary condition is imposed everywhere and there are discontinuities at the two corners of the wall,* $x = 1,\ y = \pm 1$.

We use the Q_1 finite element method to discretize the convection-diffusion equation. First, we consider a moderate value for the viscosity parameter $\nu = {}^1/_{200}$.

According to Proposition 3.8, the preconditioned spectrum is contained in a circle centered at $(1,0)$ and the radius of this circle is directly related to the approximation error ε introduced by the model order reduction for SSS matrix computations. In Section 5, we show that ε can be made arbitrarily small by setting the bound of the

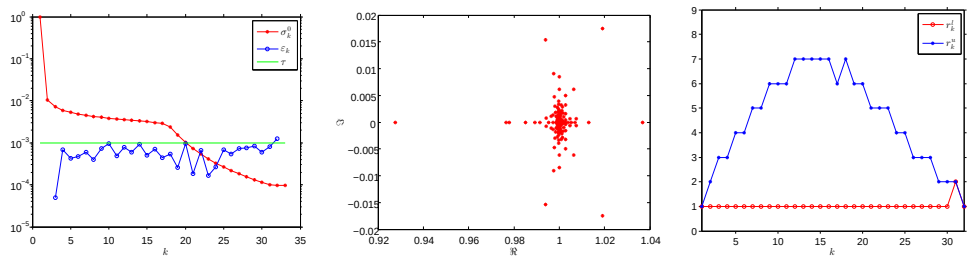
*IFISS is a computational laboratory for experimenting with state-of-the-art preconditioned iterative solvers for the discrete linear equations that arise in incompressible flow modeling, which can be run under Matlab or Octave.

discarded singular values τ properly. We give detailed information of the spectrum of the preconditioned matrix of dimension 1089×1089, which corresponds to a mesh size $h = 2^{-4}$. For different values of τ, the preconditioned spectrum and the adaptive semiseparable order are plotted in Figure 7-8.

Figure 7(a) and Figure 8(a) illustrate that the error introduced by the model order reduction at step k in computing the MSSS preconditioner, which is denoted by ε_k and measured by the matrix 2-norm, is of the same order as τ. Here ε_k is computed by

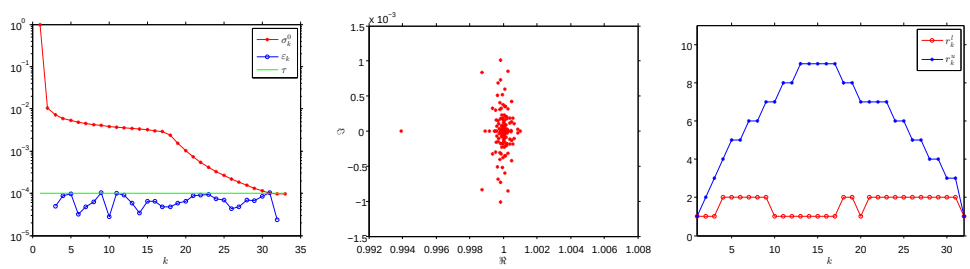
$$\varepsilon_k = \left\| \tilde{S}_k - \left(K_{k,k} - K_{k,k-1} \tilde{S}_{k-1}^{-1} K_{k-1,k} \right) \right\|_2 . \tag{6.2}$$

It also illustrates that by setting the approximation error of the model order reduction with the same order as the smallest singular value σ_0, we can compute a nonsingular preconditioner and get satisfactory convergence.



(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 7. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-3}$



(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 8. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-4}$

By decreasing τ, we get a smaller approximation error, which corresponds to smaller ε. According to our analysis, the circle that contains the preconditioned spectrum is even smaller. This is depicted by Figure 7(b)-Figure 8(b). Moreover, it is shown that by decreasing τ by a factor of 10, the radius of the circle that contains the preconditioned spectrum also decreases by a factor of about 10. This validates our bound of the radius in Proposition 3.8. In fact, for the preconditioned spectrum in Figure 7(b), only 4 iterations are needed to compute the solution by the `IDR(4)` solver and only 2 iterations are needed to solve the system corresponding to the preconditioned spectrum in Figure 8(b).

Figure 7(c)-Figure 8(c) give the maximum adaptive rank for the off-diagonal blocks of the Schur complement at step k to compute the MSSS preconditioner. Since we have an unsymmetric matrix, the τ-rank for the upper-triangular part and the

lower-triangular part are different from each other. Here the τ-rank represents the number of singular values that is bigger than or equal to τ for a matrix. Figure 7(c)-Figure 8(c) illustrate that the upper semiseparable order r^u is bigger than the lower semiseparable order r^l which states that the upper-triangular part is more difficult to approximate. Both r^l and r^u are small and this gives small average semiseparable order, which yields linear computational complexity.

We plot the spectrum of the system without preconditioning in Figure 9 to compare with the preconditioned spectrum in Figure 7(b)-Figure 8(b).

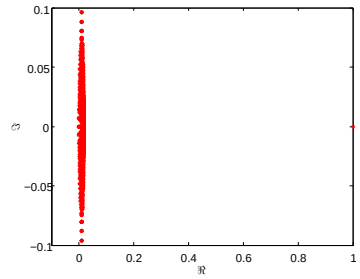
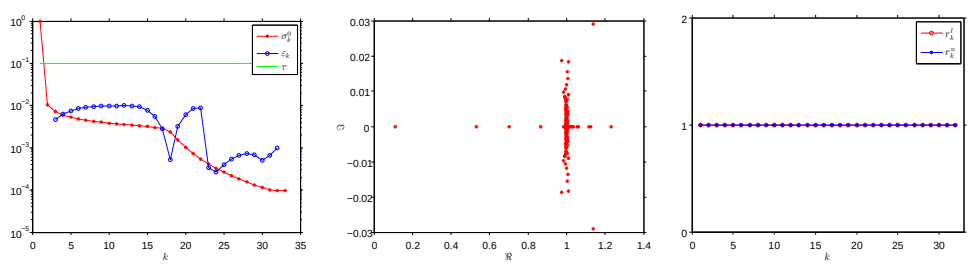


FIG. 9. *Spectrum of the system without preconditioning*

The driving force of preconditioning is to push the eigenvalues away from 0 and make them cluster. We have already seen that moderate or small setting of the model reduction error give satisfactory results. Next, we use a big model order reduction error bound by setting $\tau = 10^{-1}$ and test the performance of the MSSS preconditioner. The approximation error at each step in computing the MSSS preconditioner is plotted in Figure 10(a), the preconditioned spectrum is given in Figure 10(b), and the adaptive semiseparable order is given in Figure 10(c).

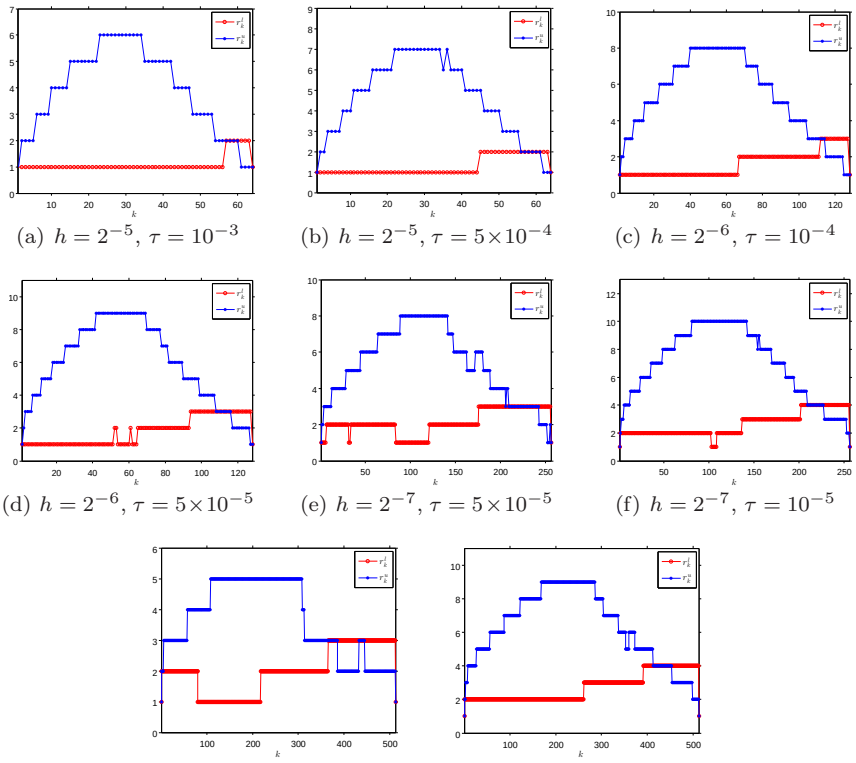


(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 10. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-1}$

Note that even this setting of the error bound is much bigger than the smallest singular value of the leading principle sub-matrices, which is used to guarantee to compute a nonsingular preconditioner, we still compute an MSSS preconditioner. This is because the possibility of perturbing a nonsingular matrix to singularity is quite small. Since we have a preconditioner that is less accurate because we use a relatively big error bound, the radius of the circle that contains the spectrum of the preconditioned matrix in Figure 10(b) is not as small as the radius in Figure 7(b)-Figure 8(b). However, the spectrum is away from 0 and only a few eigenvalues are out of this cluster. `IDR(4)` computes the solution in just 8 iterations. Moreover, the semiseparable order for such computations is just 1 as shown in Figure 10(c), which makes the computational complexity even smaller.

The performance of MSSS preconditioners for different mesh sizes h and τ are reported in Table 1. For different settings of τ, the adaptive semiseparable order is given in Figure 11.



(a) $h = 2^{-5}$, $\tau = 10^{-3}$ (b) $h = 2^{-5}$, $\tau = 5\times10^{-4}$ (c) $h = 2^{-6}$, $\tau = 10^{-4}$

(d) $h = 2^{-6}$, $\tau = 5\times10^{-5}$ (e) $h = 2^{-7}$, $\tau = 5\times10^{-5}$ (f) $h = 2^{-7}$, $\tau = 10^{-5}$

(g) $h = 2^{-8}$, $\tau = 5\times10^{-5}$ (h) $h = 2^{-8}$, $\tau = 10^{-5}$

FIG. 11. *Adaptive semiseparable order for convection-diffusion equation with* $\nu = {}^{1}/_{200}$

TABLE 1

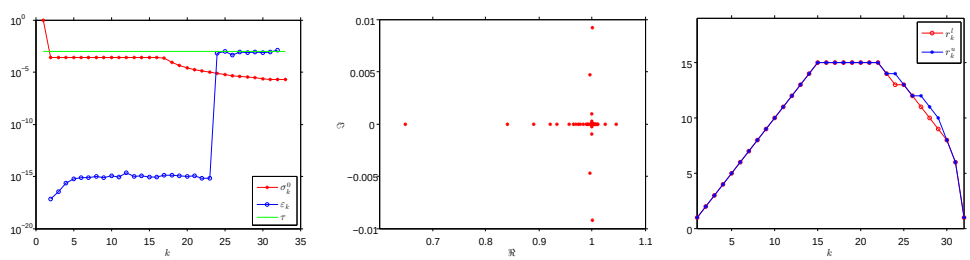
Computational results of the MSSS preconditioner for the convection-diffusion equation with $\nu = {}^{1}/_{200}$

h	N^2	τ	# iter.
2^{-4}	$1.09e+03$	5×10^{-3}	6
		10^{-3}	3
2^{-5}	$4.23e+03$	10^{-3}	7
		5×10^{-4}	4
2^{-6}	$1.66e+04$	10^{-4}	5
		5×10^{-5}	4
2^{-7}	$6.61e+04$	5×10^{-5}	6
		10^{-5}	3
2^{-8}	$2.63e+05$	5×10^{-5}	10
		10^{-5}	5

The results reported in Table 1 and Figure 11 state that by choosing a proper error bound for the model order reduction, we can compute an MSSS preconditioner that gives satisfactory convergence. This convergence can be made independent of the mesh size, while the adaptive semiseparable order only slightly increases with the problem size. This is demonstrated by Figure 11. The average of the upper and lower semiseparable order is still quite small and can be almost kept constant. We can also choose a fixed semiseparable order to compute an MSSS preconditioner. This is studied in [32]. Both ways to compute the MSSS preconditioner give satisfactory results.

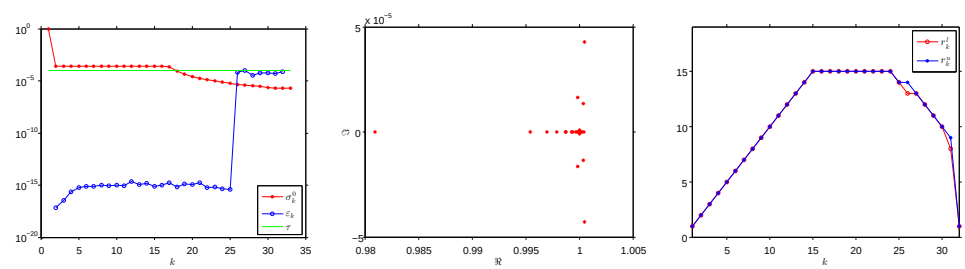
Next we set $\nu = 10^{-4}$ for the convection-diffusion equation, which corresponds to the convection-dominated case. For such test case, the finite element discretization is not stable anymore, an up-wind scheme should be applied to get a stable discretization. Due to the ill-conditioning of the system, a bigger semiseparable order is needed to compute the MSSS preconditioner to get better performance. Note that for this case, the multigrid methods (both AMG and GMG) fail to solve such system without up-wind scheme while the MSSS preconditioner can still solve this ill-conditioning systems [32]. Here we report detailed numerical results for the test case with mesh size $h = 2^{-4}$. We first set $\tau = 10^{-3}$ and $\tau = 10^{-4}$, the computational results are reported in Figure 12 and Figure 13. The preconditioned system can be solved by `IDR(4)` using 4 iterations for $\tau = 10^{-3}$ and 2 iterations for $\tau = 10^{-4}$.

Figure 12(b)-Figure 13(b) show that by reducing τ with a factor of 10, the radius of the circle that contains the preconditioned spectrum is also reduced by a factor around 10. This validates our bound of the radius of circle that contains the preconditioned spectrum in Proposition 3.8 again.



(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 12. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-3}$



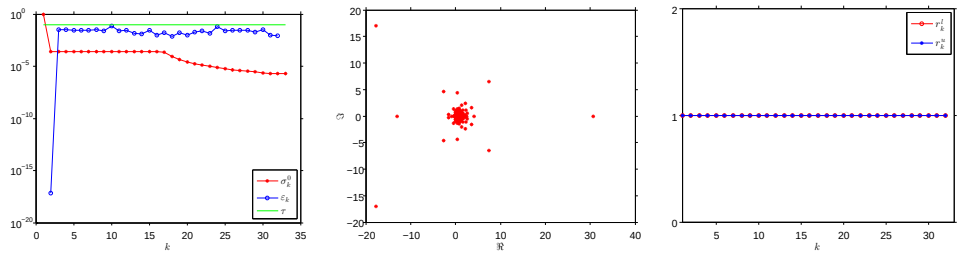
(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 13. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-4}$

REMARK 6.1. The MSSS preconditioners computed by setting $\tau = 10^{-3}$ and $\tau = 10^{-4}$ give different clustering of the spectrum, as shown in Figure 12(b)-Figure 13(b).

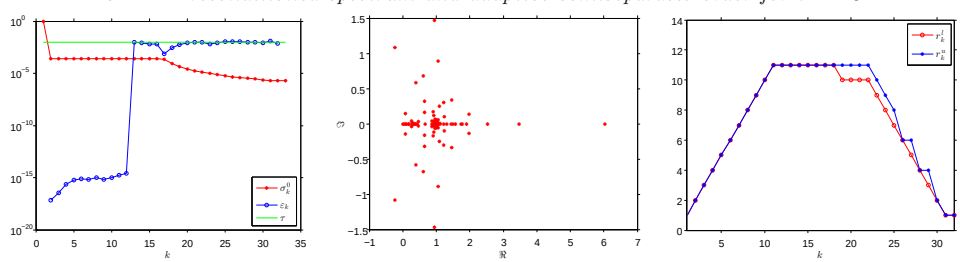
However, the adaptive semiseparable order for different τ is almost the same. This is primarily because the Schur complements in computing the factorization are very ill-conditioned and difficult to approximate. A slight change of the semiseparable order results in relatively big difference of the approximation accuracy. This also explains why a bigger adaptive semiseparable order is needed compared with the test case for $\nu = {}^{1}\!/_{200}$.

We also test the convergence of the MSSS preconditioned system by setting $\tau = 10^{-1}$ and $\tau = 10^{-2}$. The computational results are given in Figure 14-Figure 15.



(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 14. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-1}$



(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 15. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-2}$

It is shown in Figure 14-Figure 15 that although the MSSS preconditioner is non-singular by choosing relatively bigger τ compared to σ_0, the preconditioned spectrum is contained in a circle centered at $(1, 0)$ with a much bigger radius and $(0, 0)$ is included in the circle. The preconditioned system for $\tau = 10^{-2}$ is solved by `IDR(4)` in 61 iterations while `IDR(4)` fails to solve the preconditioned system for $\tau = 10^{-1}$ within 80 iterations.

To give an impression of how difficult it is to preconditioning this convection-dominated system, we plot part of the spectrum in Figure 16. It is shown that a big part of the eigenvalues is close to 0, which makes slow convergence.

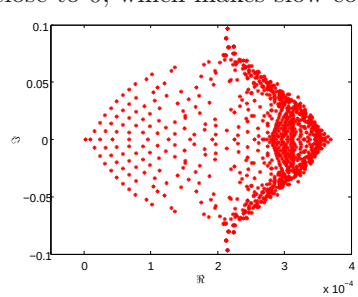
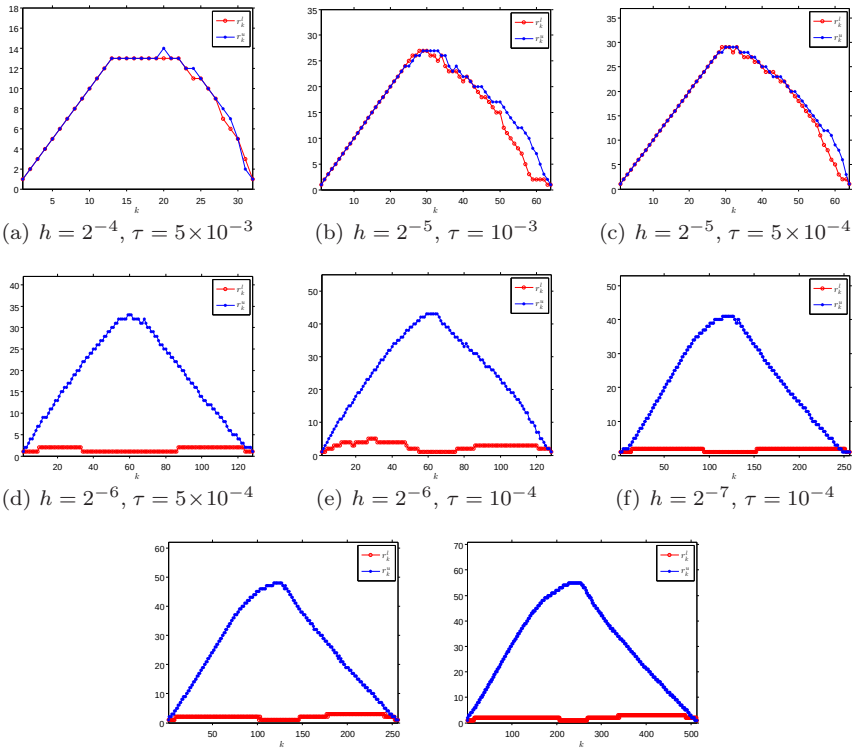


FIG. 16. *Part of the non-preconditioned spectrum*

For different mesh sizes h and settings of τ, the computational results are reported in Table 2. The adaptive semiseparable order for different mesh sizes h and settings of τ are plotted in Figure 17.

TABLE 2
Computational results of the MSSS preconditioner for the convection-diffusion equation with $\nu = 10^{-4}$

h	N^2	τ	# iter.
2^{-4}	$1.09e+03$	5×10^{-3}	23
		10^{-3}	4
2^{-5}	$4.23e+03$	10^{-3}	18
		5×10^{-4}	11
2^{-6}	$1.66e+04$	5×10^{-4}	20
		10^{-4}	6
2^{-7}	$6.61e+04$	10^{-4}	13
		5×10^{-5}	8
2^{-8}	$2.63e+05$	5×10^{-5}	17
		10^{-5}	6



(a) $h = 2^{-4}$, $\tau = 5\times10^{-3}$ (b) $h = 2^{-5}$, $\tau = 10^{-3}$ (c) $h = 2^{-5}$, $\tau = 5\times10^{-4}$

(d) $h = 2^{-6}$, $\tau = 5\times10^{-4}$ (e) $h = 2^{-6}$, $\tau = 10^{-4}$ (f) $h = 2^{-7}$, $\tau = 10^{-4}$

(g) $h = 2^{-7}$, $\tau = 5\times10^{-5}$ (h) $h = 2^{-8}$, $\tau = 10^{-5}$

FIG. 17. *Adaptive semiseparable order for convection-diffusion equation with* $\nu = 10^{-4}$

Since this convection-dominated test case is ill-conditioned, it is quite difficult to compute its factorization (inverse). It takes more effort to compute a good ap-

proximation compared with the case $\nu = {}^1/_{200}$. This is illustrated by comparing the adaptive semiseparable order in Figure 17 with that in Figure 11. For such case, the average of the upper and lower semiseparable order is considerably bigger. Since the average semiseparable order may not be bounded by a small constant, this makes the computational complexity of the MSSS preconditioning technique slightly bigger than linear. Details and remarks on the computational complexity for moderate semiseparable order will be discussed in Section 6.2.

6.2. Symmetric Indefinite Systems from Discretized Helmholtz Equation. In this subsection, we study the convergence of the MSSS preconditioners for the symmetric indefinite systems from the discretization of scalar PDEs, where the Schrödinger equation in Example 4.1 and the Helmholtz equation belong to this type. In this part, we mainly focus on the performance of the MSSS preconditioner for the Helmholtz equation that is given by Example 6.2.

EXAMPLE 6.2 ([25]). *Consider the following Helmholtz equation,*

$$-\nabla^2 u(x,\omega) - \frac{\omega^2}{c^2(x)} u(x,\omega) = g(x,\omega),\ \ x \in [0,1] \times [0,1],$$

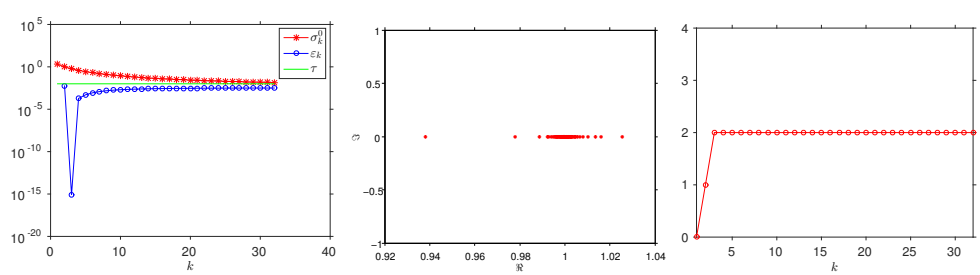
with homogeneous Dirichlet boundary condition. Here $u(x,\omega)$ represents the pressure field in the frequency domain, ∇^2 is the Laplacian operator, ω is the angular frequency, and $c(x)$, is the acoustic-wave velocity that varies with position x.

Standard five-point stencil finite difference method is used to discretize the Helmoltz equation. We use the rule of thumb that at least 10 nodes per wavelength should be employed, which leads to the restriction

$$\kappa h \leq \frac{\pi}{5} \approx 0.628, \tag{6.3}$$

for the standard five-point stencil finite difference discretization [25]. Here $\kappa = {}^{\omega}/_{c(x)}$ is the wave number. We apply the MSSS preconditioner to the Helmholtz equation. The pulse source term $g(x,\omega)$ is chosen as the scaled delta function that is located at $\left({}^1/_{32}, {}^1/_2\right)$.

To test the performance of the MSSS preconditioner, we first set a moderate value for kh, say ${}^1/_{16}$. The preconditioned spectrum and semiseparable order for mesh size $h = 2^{-5}$ and different settings of τ are plotted in Figure 18 and Figure 19.

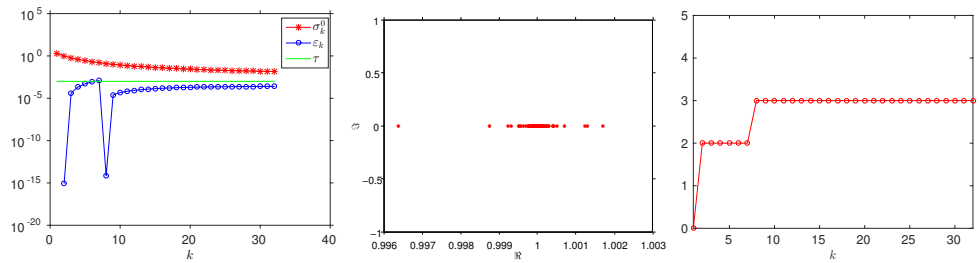


(a) Minimal singular value and ε_k, $\kappa = 2$ (b) Preconditioned spectrum, $\kappa = 2$ (c) Adaptive semiseparable order, $\kappa = 2$

FIG. 18. *Preconditioned spectrum and adaptive semiseparable order for $\tau = 10^{-2}$ and $\kappa h = {}^1/_{16}$*

For $\tau = 10^{-2}$, the error introduced by the model order reduction is already smaller than σ_0, we confer that the preconditioned spectrum is contained in a circle that is small enough according to Proposition 3.8. This is well illustrated by Figure 18(b). If

we reduce τ to 10^{-3}, then we get even smaller circle that contains the preconditioned spectrum in Figure 19(b). Figure 18(b)-Figure 19(b) indicate that by reducing τ with a factor of 10, the radius of the circle that contains the preconditioned spectrum also decreases by a factor around 10. This again verifies our analysis on the radius of the circle in Proposition 3.8.



(a) Minimal singular value and ε_k, $\kappa = 2$ (b) Preconditioned spectrum, $\kappa = 2$ (c) Adaptive semiseparable order, $\kappa = 2$

FIG. 19. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-3}$ *and* $\kappa h = {}^1/_{16}$

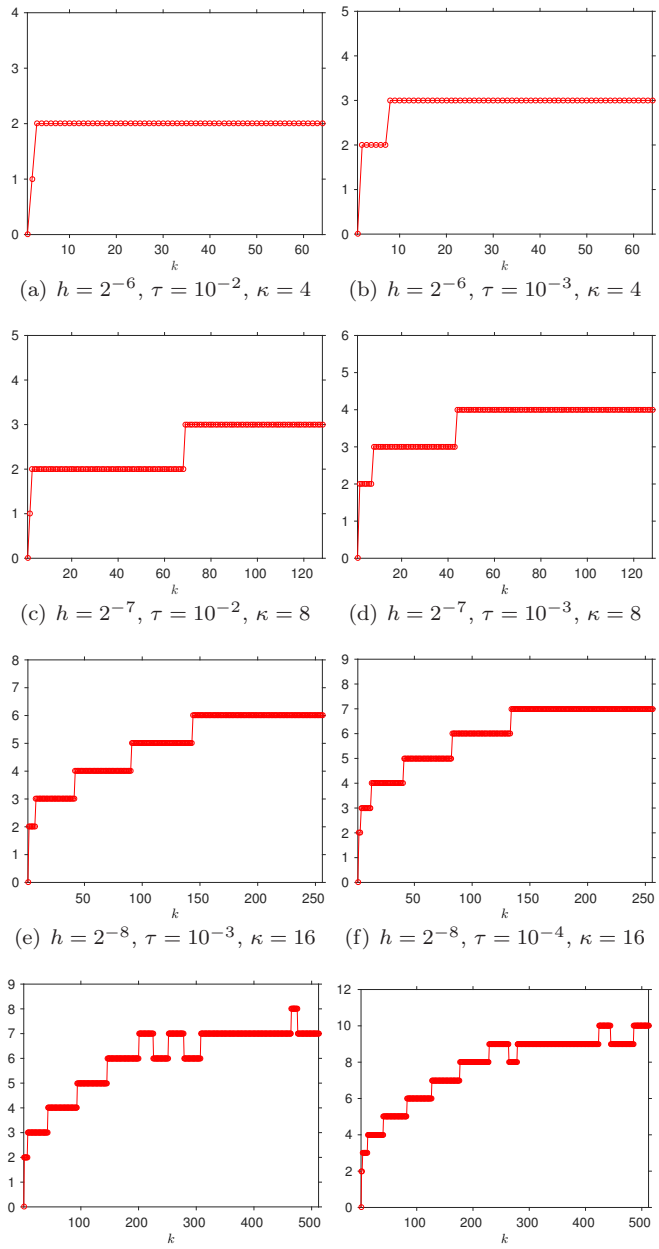
Both settings of τ give small enough circle to yield very fast convergence for Krylov solvers. `IDR(4)` computes the solution in 4 iterations for $\tau = 10^{-2}$ and 2 iterations for $\tau = 10^{-3}$. For both cases, the semiseparable order is small enough to yield linear computational complexity of the MSSS preconditioners.

For the settings of different τ and h, the computational results are reported in Table 3. The adaptive semiseparable order are plotted in Figure 20.

TABLE 3
Performance of MSSS preconditioner for the Helmholtz equation with $\kappa h = {}^1/_{16}$

h	κ	N^2	τ	# iter.
2^{-5}	2	$1.09e+03$	10^{-2}	4
			10^{-3}	2
2^{-6}	4	$4.23e+03$	10^{-2}	6
			10^{-3}	3
2^{-7}	8	$1.66e+04$	10^{-2}	8
			10^{-3}	4
2^{-8}	16	$6.61e+04$	10^{-3}	7
			10^{-4}	3
2^{-9}	32	$2.63e+05$	10^{-3}	14
			10^{-4}	3

The computational results in Table 3 show that the number of iterations can be kept constant by setting a proper τ. We keep κh constant for numerical experiments. This makes the Helmholtz equation even more difficult to solve, and results in a slight increase of the semiseparable order for big κ in Figure 20. However, the semiseparable order is still bounded by a small number.



(a) $h = 2^{-6}$, $\tau = 10^{-2}$, $\kappa = 4$ (b) $h = 2^{-6}$, $\tau = 10^{-3}$, $\kappa = 4$

(c) $h = 2^{-7}$, $\tau = 10^{-2}$, $\kappa = 8$ (d) $h = 2^{-7}$, $\tau = 10^{-3}$, $\kappa = 8$

(e) $h = 2^{-8}$, $\tau = 10^{-3}$, $\kappa = 16$ (f) $h = 2^{-8}$, $\tau = 10^{-4}$, $\kappa = 16$

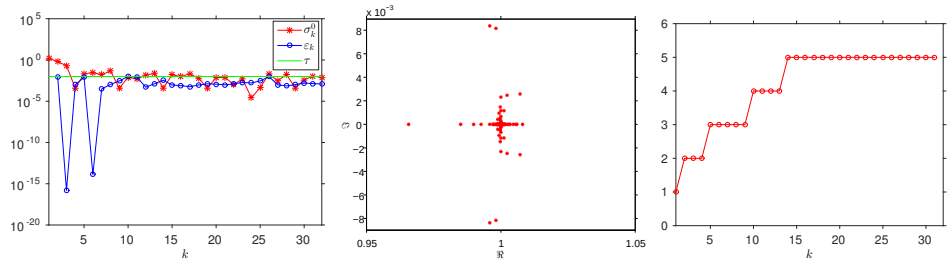
(g) $h = 2^{-9}$, $\tau = 10^{-3}$, $\kappa = 32$ (h) $h = 2^{-9}$, $\tau = 10^{-4}$, $\kappa = 32$

FIG. 20. *Adaptive semiseparable order for the Helmholtz equation with* $kh = {}^{1}/_{16}$

Next, we set $\kappa h = 0.625$ that almost reaches the limit of condition (6.3). We first report the preconditioned spectrum to demonstrate our analysis. By choosing different settings of τ, the computational results for the mesh size $h = 2^{-5}$ are given in Figure 21-Figure 22.

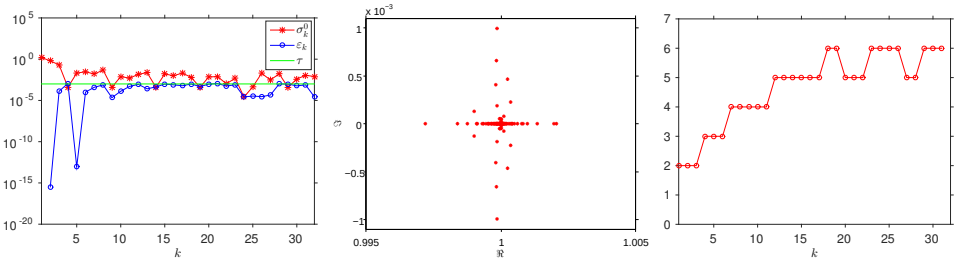
τ is first chosen as 10^{-2}, which is of $\mathcal{O}(\sigma_0)$. This gives the computational results in Figure 21. The preconditioned spectrum in Figure 21(b) is contained in a circle with very small radius. By decreasing τ, we get a even smaller radius of the circle in Figure 22(b). It is shown in Figure 21(b)-Figure 22(b) that if τ is decreased by

a factor of 10, the radius of the circle that contains the preconditioned spectrum is also reduced by a factor around 10. Both settings give super fast convergence. Here the Helmholtz problem is solved by `IDR(4)` in only 3 iterations for $\tau = 10^{-2}$ and 2 iterations for $\tau = 10^{-3}$.



(a) Minimal singular value and ε_k, $\kappa = 20$ (b) Preconditioned spectrum, $\kappa = 20$ (c) Adaptive semiseparable order, $\kappa = 20$

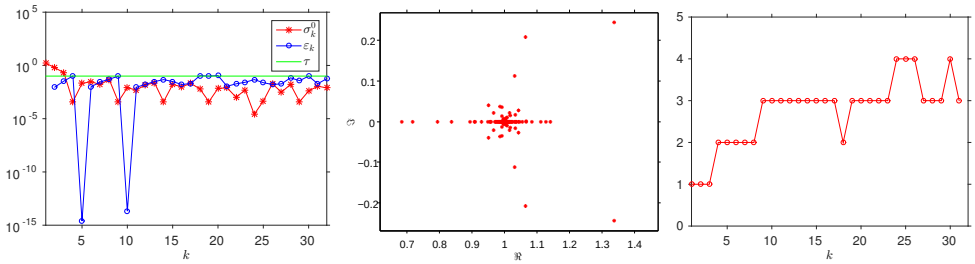
FIG. 21. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-2}$



(a) Minimal singular value and ε_k, $\kappa = 20$ (b) Preconditioned spectrum, $\kappa = 20$ (c) Adaptive semiseparable order, $\kappa = 20$

FIG. 22. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-3}$

We can even relax the settings of τ and still compute an efficient MSSS preconditioner. This setting gives us smaller semiseparable order that is preferable. The computational results are shown in Figure 23. The preconditioned spectrum is contained in a circle with a bigger radius compared with the case $\tau = 10^{-2}$. Compare Figure 23(b) with Figure 22(b), we see that by increasing τ with a factor of 10, the radius of the circle that contains the preconditioned spectrum also increases by a factor around 10. This is stated in Proposition 3.8. The circle that contains the preconditioned spectrum still has a small radius for $\tau = 10^{-1}$. Therefore, `IDR(4)` computes the solution in only 9 iterations.

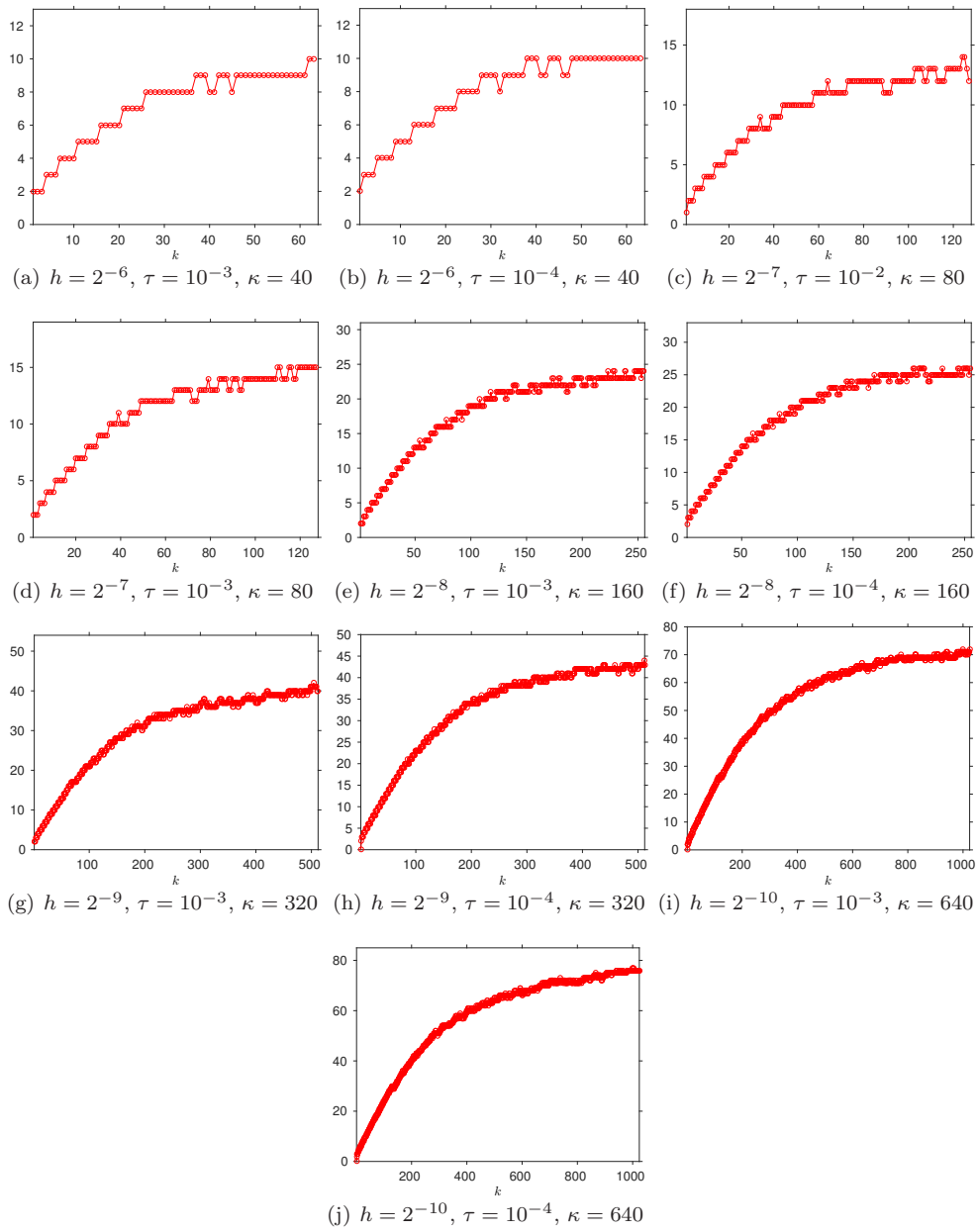


(a) Minimal singular value and ε_k, $\kappa = 20$ (b) Preconditioned spectrum, $\kappa = 20$ (c) Adaptive semiseparable order, $\kappa = 20$

FIG. 23. *Preconditioned spectrum and adaptive semiseparable order for* $\tau = 10^{-1}$

We report the computational results of the MSSS preconditioner for different

mesh sizes h and settings of τ in Table 4. The adaptive semiseparable order for the MSSS preconditioners are plotted in Figure 24.



(a) $h = 2^{-6}$, $\tau = 10^{-3}$, $\kappa = 40$ (b) $h = 2^{-6}$, $\tau = 10^{-4}$, $\kappa = 40$ (c) $h = 2^{-7}$, $\tau = 10^{-2}$, $\kappa = 80$

(d) $h = 2^{-7}$, $\tau = 10^{-3}$, $\kappa = 80$ (e) $h = 2^{-8}$, $\tau = 10^{-3}$, $\kappa = 160$ (f) $h = 2^{-8}$, $\tau = 10^{-4}$, $\kappa = 160$

(g) $h = 2^{-9}$, $\tau = 10^{-3}$, $\kappa = 320$ (h) $h = 2^{-9}$, $\tau = 10^{-4}$, $\kappa = 320$ (i) $h = 2^{-10}$, $\tau = 10^{-3}$, $\kappa = 640$

(j) $h = 2^{-10}$, $\tau = 10^{-4}$, $\kappa = 640$

FIG. 24. *Adaptive semiseparable order for the Helmholtz equation with* $\kappa h = 0.625$

Results listed in Table 4 illustrate that by properly setting τ, we obtain mesh size h independent and wave number k independent convergence. The number of iterations can be kept virtually constant. For the shifted Laplacian preconditioner [17], which is the state-of-the-art preconditioning technique for the Helmholtz equation, the number of iterations scales linearly with the wave number κ [43]. Some recent effort dedicated to reduce the dependency on wave number of the number of iterations is carried out by making use of the deflation technique, cf. [37].

With the refinement of the mesh, the wave number κ increases linearly. This makes the Helmholtz problem difficult to solve for small mesh size, which is illustrated by the increase of semiseparable order. Figure 24 shows that the semiseparable order has a considerable increase with the refinement of the mesh and is not bounded by a small number, but a moderate number of $\mathcal{O}(\sqrt{N})$. As stated in [8] that the computational complexity for the SSS matrix computations is linear with respect to the problem size with a prefactor r_k^3 provided that r_k is small. Next, we analyze the computational complexity of SSS matrix computations for big r_k but $r_k \ll N$, which corresponds to the case of the Helmholtz equation for $\kappa h = 0.625$.

Table 4
Performance of MSSS preconditioner for the Helmholtz equation with $\kappa h = 0.625$

h	κ	N^2	τ	# iter.
2^{-5}	20	$1.09e+03$	10^{-2}	3
			10^{-3}	2
2^{-6}	40	$4.23e+03$	10^{-3}	3
			10^{-4}	3
2^{-7}	80	$1.66e+04$	10^{-2}	8
			10^{-3}	4
2^{-8}	160	$6.61e+04$	10^{-3}	5
			10^{-4}	3
2^{-9}	320	$2.63e+05$	10^{-3}	5
			10^{-4}	3
2^{-10}	640	$1.05e+06$	10^{-3}	6
			10^{-4}	3

For an SSS matrix A of $N \times N$ with semiseparable order r_k, the size of its diagonal blocks is denoted by n, then A has $^N/_n$ blocks. The computational complexity for the matrix-matrix operations and the model order reduction of SSS matrices is bounded by

$$\mathcal{O}(\max\left\{n^3,\ n^2 r_k,\ r_k^2 n,\ r_k^3\right\} \frac{N}{n}), \tag{6.4}$$

which can be obtained by checking the SSS matrix computations in [8]. For small r_k, we also set n small enough. This gives the computational complexity of $\mathcal{O}(r_k^3 N)$, i.e., linear with respect to the problem size. For moderate r_k, the term r_k^3 becomes big. According to (6.4), the settings of n can be adjusted such that a proper computational complexity can be reached. Usually, we choose n and r_k of the same order, say $r_k = n$. This in turn gives the computational complexity for SSS matrix computations of $\mathcal{O}(r_k^2 N)$ for moderate r_k. Note that the setting of n does not change r_k, since r_k is the rank of the off-diagonal blocks, which only depends on the property of the matrix.

For the computations of the MSSS preconditioners with moderate r_k, N Schur complements are computed while each Schur complement is computed in $\mathcal{O}(r_k^2 N)$. This in turn gives the total computational complexity $\mathcal{O}(r_k^2 N^2)$, where N^2 is the problem size. For the case $\kappa h = 0.625$, r_k is roughly bounded by $\mathcal{O}(\sqrt{N})$, this in turn results the total computational complexity bounded by $\mathcal{O}(N^2)^{\frac{3}{2}}$. This is comparable

with the computational complexity of a multi-frontal solver for 2D problems [16]. We use Figure 25 to show the growth factor of time to compute the MSSS preconditioner with the mesh refinement.

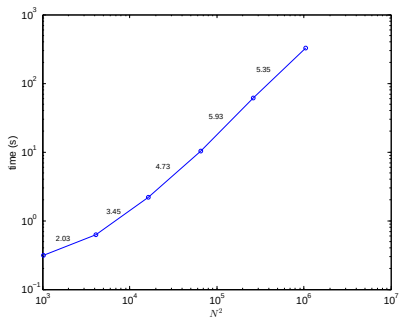


FIG. 25. *MSSS preconditioning time for* $\kappa h = 0.625$

Here we set $n = 4,\ 8,\ 8,\ 16,\ 32,\ 128$ with the refinement of the mesh and $\tau = 10^{-3}$ for all the mesh sizes except $\tau = 10^{-4}$ for $h = 2^{-9}$. All the time are measured in seconds. The number over the line shows the growth factor of the time to compute the MSSS preconditioner. It is clear that the growth factor for time is below $4^{\frac{3}{2}} = 8$. Note that we use a non-equidistant axis for Figure 25.

6.3. Saddle-Point Systems. We study the convergence property of MSSS preconditioners for the saddle-point systems in this part. Consider the following PDE-constrained optimization problem given by Example 6.3.

EXAMPLE 6.3 ([29]). *Let* $\Omega = [0,\ 1]^2$ *and consider the problem*

$$\begin{aligned} \min_{u,f} \frac{1}{2}\|u - \hat{u}\| + \frac{\beta}{2}\|f\|^2 \\ s.t.\ -\nabla^2 u = f \ \ in\ \Omega \\ u = u_D \ \ on\ \Gamma_D, \end{aligned}$$

where $\Gamma_D = \partial\Omega$, $\beta > 0$, $\hat{u} = 0$ *is the prescribed system state, and*

$$u_D = \begin{cases} -\sin(2\pi y) & if\ \ x = 1, 0 \le y \le 1, \\ \sin(2\pi y) & if\ \ x = 0, 0 \le y \le 1, \\ 0 & otherwise. \end{cases}$$

Discretize the cost function and the PDE constraints by using the Galerkin method and then compute the optimality condition gives the following linear saddle-point system to solve

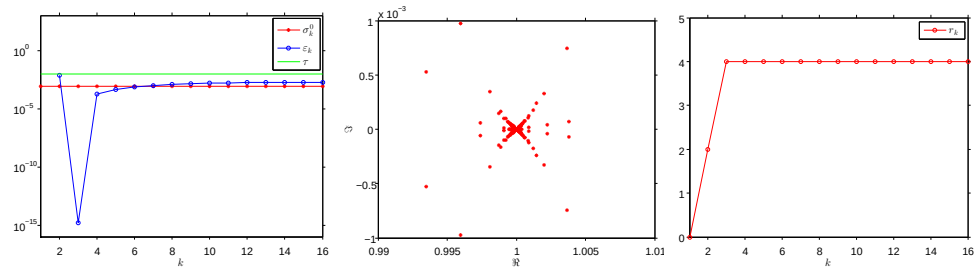
$$\begin{bmatrix} 2\beta M & 0 & -M \\ 0 & M & K^T \\ -M & K & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ b \\ d \end{bmatrix}. \tag{6.5}$$

Here M is the mass matrix, K is the stiffness matrix, x and y are the discrete analog of f and u, λ is the Lagrangian multiplier, b and d are obtained by discretizing the cost function and boundary conditions, respectively.

All the sub-blocks of the saddle-point system (6.5) have an MSSS structure and can be exploited to get a global MSSS structure. By exploiting the global MSSS structure of the saddle-point system, we can compute a global MSSS preconditioner.

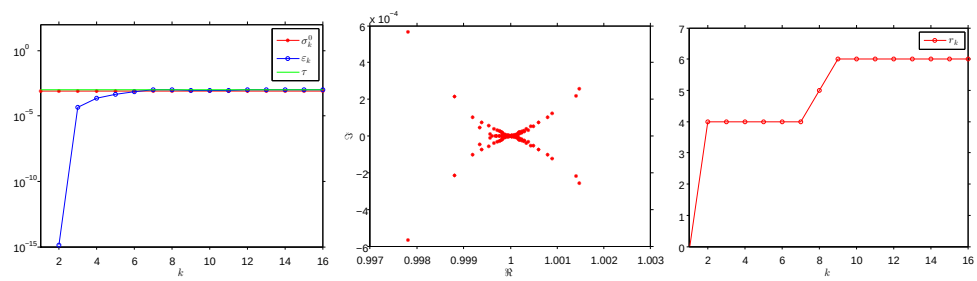
This is discussed in great detail in [31]. Here, we use the numerical experiments of preconditioning the saddle-point system (6.5) to demonstrate the convergence analysis in Section 3.

We report the computational results for the mesh size $h = 2^{-4}$ with a wide range settings of β and τ. First, we test a moderate set of $\tau = 10^{-2}$ and $\tau = 10^{-3}$ for $\beta = 10^{-1}$. The computational results are given in Figure 26 and Figure 27.



(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 26. *Preconditioned spectrum and adaptive semiseparable order for* $h = 2^{-4}$, $\tau = 10^{-2}$, $\beta = 10^{-1}$



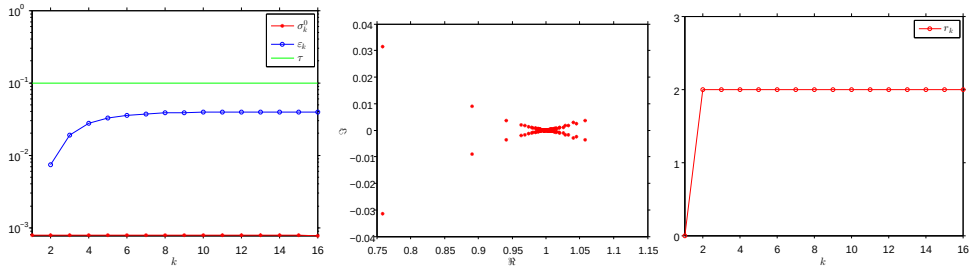
(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 27. *Preconditioned spectrum and adaptive semiseparable order for* $h = 2^{-4}$, $\tau = 10^{-3}$, $\beta = 10^{-1}$

Figure 26(b) and Figure 27(b) show that a moderate setting of τ gives a small radius of the circle that contains the preconditioned spectrum, and the smaller τ is, the smaller the radius is. Both settings give adequately small circle and the decrease of τ just yields a slightly increase of the semiseparable order, which is shown in Figure 26(c) and Figure 27(c). The `IDR(4)` solves the preconditioned system for $\tau = 10^{-2}$ in only 3 iterations and only 2 iterations for $\tau = 10^{-3}$.

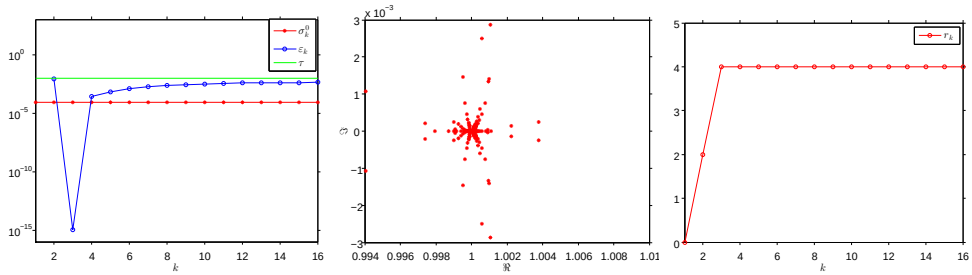
We can even set a bigger τ to compute an MSSS preconditioner, the computational results for $\tau = 10^{-1}$ are plotted in Figure 28. Figure 28(b) illustrates that a bigger τ gives a bigger radius of the circle that contains the preconditioned spectrum. But the circle is still small and the `IDR(4)` solver computes the solution of the preconditioned system in only 5 iterations. Moreover, this settings of τ gives a smaller semiseparable order, which is shown in Figure 28(c).

The smallest singular value of the saddle-point system (6.5) scales with β. A smaller β in turn gives a smaller smallest singular value. This makes the saddle-point system (6.5) even more ill-conditioned and difficult to solve. Next, we test the case for a moderate $\beta = 10^{-2}$, and a much smaller $\beta = 10^{-5}$, the computational results for $\tau = 10^{-2}$, and $\tau = 10^{-3}$ are reported in Figure 29-Figure 32.



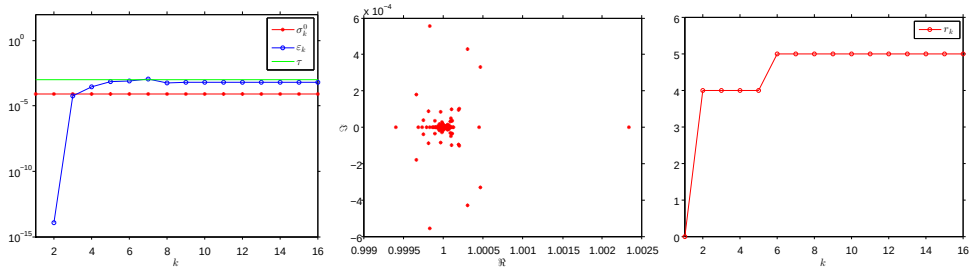
(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 28. *Preconditioned spectrum and adaptive semiseparable order for* $h = 2^{-4}$, $\tau = 10^{-1}$, $\beta = 10^{-1}$



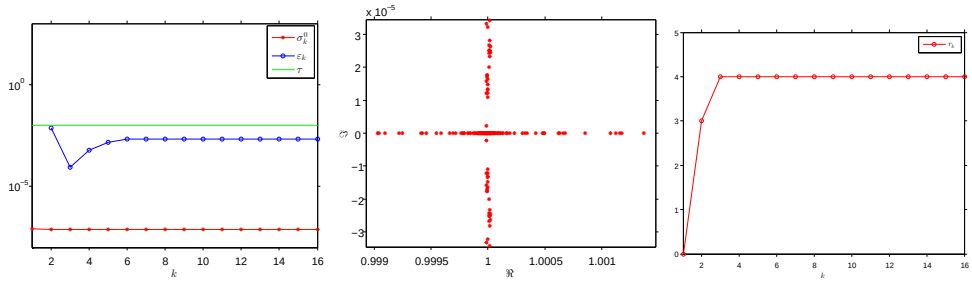
(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 29. *Preconditioned spectrum and adaptive semiseparable order for* $h = 2^{-4}$, $\tau = 10^{-2}$, $\beta = 10^{-2}$



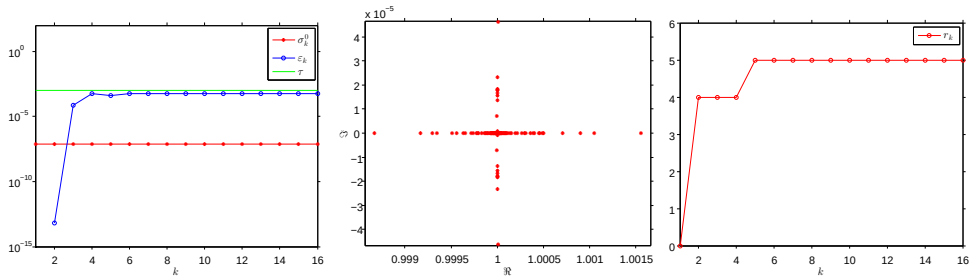
(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 30. *Preconditioned spectrum and adaptive semiseparable order for* $h = 2^{-4}$, $\tau = 10^{-3}$, $\beta = 10^{-2}$



(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 31. *Preconditioned spectrum and adaptive semiseparable order for* $h = 2^{-4}$, $\tau = 10^{-2}$, $\beta = 10^{-5}$



(a) Minimal singular value and ε_k (b) Preconditioned spectrum (c) Adaptive semiseparable order

FIG. 32. *Preconditioned spectrum and adaptive semiseparable order for* $h = 2^{-4}$, $\tau = 10^{-3}$, $\beta = 10^{-5}$

The computational results in Figure 29-Figure 32 show that a moderate settings of τ gives satisfactory performance of the MSSS preconditioner. The `IDR(4)` solver computes the solution in 2 or 3 iterations for all the settings of τ and β that corresponds to the test cases in Figure 29-Figure 32. Although the smallest singular value of all the principle leading sub-matrices of the saddle-point system (6.5) is much smaller for small β, the choice of τ can still be made bigger than the smallest singular value, which is illustrated by Figure 29(a)-Figure 32(a). The choice of moderate τ in turn gives an adequately small semiseparable order for a wide range of β.

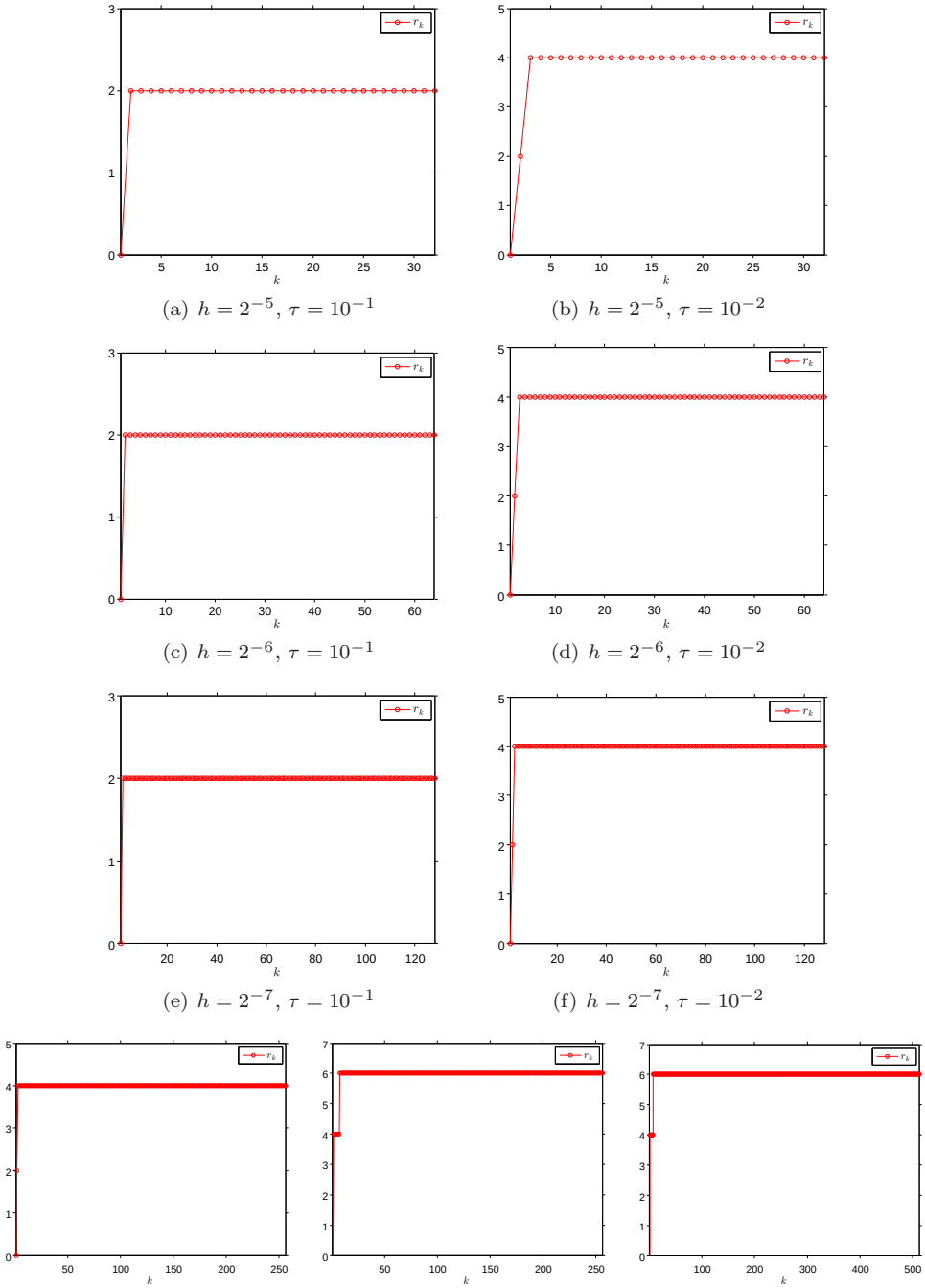
TABLE 5
Performance of MSSS preconditioner for the PDE-constrained optimization problem with different β

h	N^2	$\beta = 10^{-1}$		$\beta = 10^{-2}$		$\beta = 10^{-5}$	
		τ	# iter.	τ	# iter.	τ	# iter.
2^{-5}	$3.07e+03$	10^{-1}	6	10^{-1}	6	10^{-1}	4
		10^{-2}	3	10^{-2}	4	10^{-2}	3
2^{-6}	$1.23e+04$	10^{-1}	9	10^{-1}	8	10^{-2}	16
		10^{-2}	4	10^{-2}	4	10^{-3}	3
2^{-7}	$4.92e+04$	10^{-1}	13	10^{-1}	16	10^{-3}	6
		10^{-2}	5	10^{-2}	5	10^{-4}	2
2^{-8}	$1.97e+05$	10^{-2}	10	10^{-2}	10	10^{-4}	3
		10^{-3}	4	10^{-3}	4		
2^{-9}	$7.86e+05$	10^{-3}	6	10^{-3}	6	10^{-4}	19
						10^{-5}	2

The performance of the MSSS preconditioner for different settings of τ, β, and the mesh size h are given in Table 5. The corresponding semiseparable order are plotted in Figure 33-Figure 35.

The computational results in Figure 33-Figure 34 for the test case with $\beta = 10^{-1}$ and $\beta = 10^{-2}$ show that for a constant setting of τ, the semiseparable order is bounded by a constant 4 for the mesh size h ranges from 2^{-7} to 2^{-5}. The semiseparable order is independent of the mesh size h and β. Since the smallest singular value for all the principle leading sub-matrices of the saddle-point systems also scales with mesh size h, for a bigger test example with mesh size $h = 2^{-9}$ and $h = 2^{-8}$, a smaller τ

is needed to get a satisfactory radius of the circle that contains the preconditioned spectrum according to Proposition 3.8. This is verified by the computational results in Figure 33-Figure 35 and Table 5. The setting of a smaller τ yields a slightly increase of the semiseparable order from 4 to 6, which is still quite small.



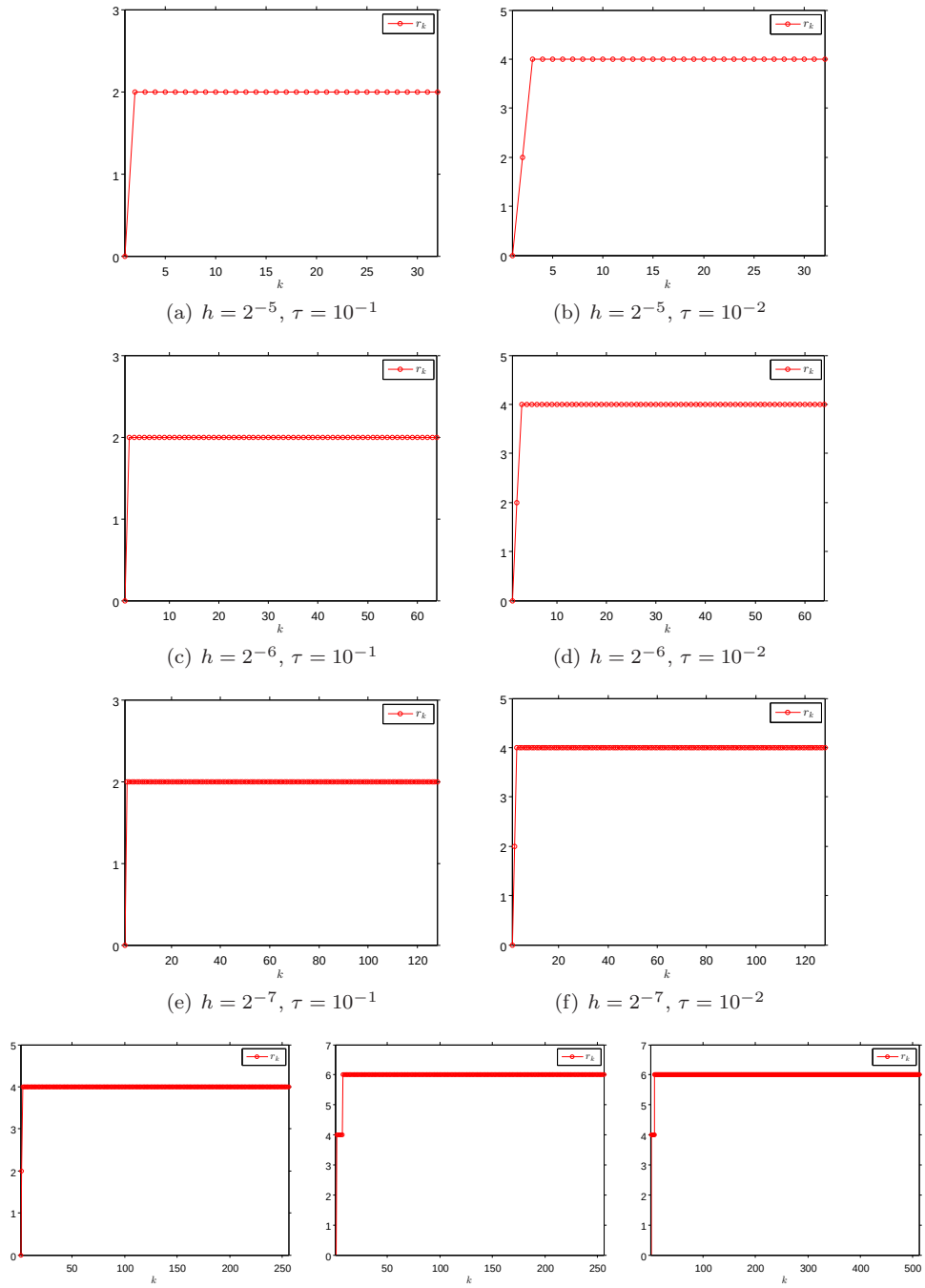
(a) $h = 2^{-5}$, $\tau = 10^{-1}$ (b) $h = 2^{-5}$, $\tau = 10^{-2}$

(c) $h = 2^{-6}$, $\tau = 10^{-1}$ (d) $h = 2^{-6}$, $\tau = 10^{-2}$

(e) $h = 2^{-7}$, $\tau = 10^{-1}$ (f) $h = 2^{-7}$, $\tau = 10^{-2}$

(g) $h = 2^{-8}$, $\tau = 10^{-2}$ (h) $h = 2^{-8}$, $\tau = 10^{-3}$ (i) $h = 2^{-9}$, $\tau = 10^{-3}$

FIG. 33. *Adaptive semiseparable order for* $\beta = 10^{-1}$



(a) $h = 2^{-5}$, $\tau = 10^{-1}$ (b) $h = 2^{-5}$, $\tau = 10^{-2}$

(c) $h = 2^{-6}$, $\tau = 10^{-1}$ (d) $h = 2^{-6}$, $\tau = 10^{-2}$

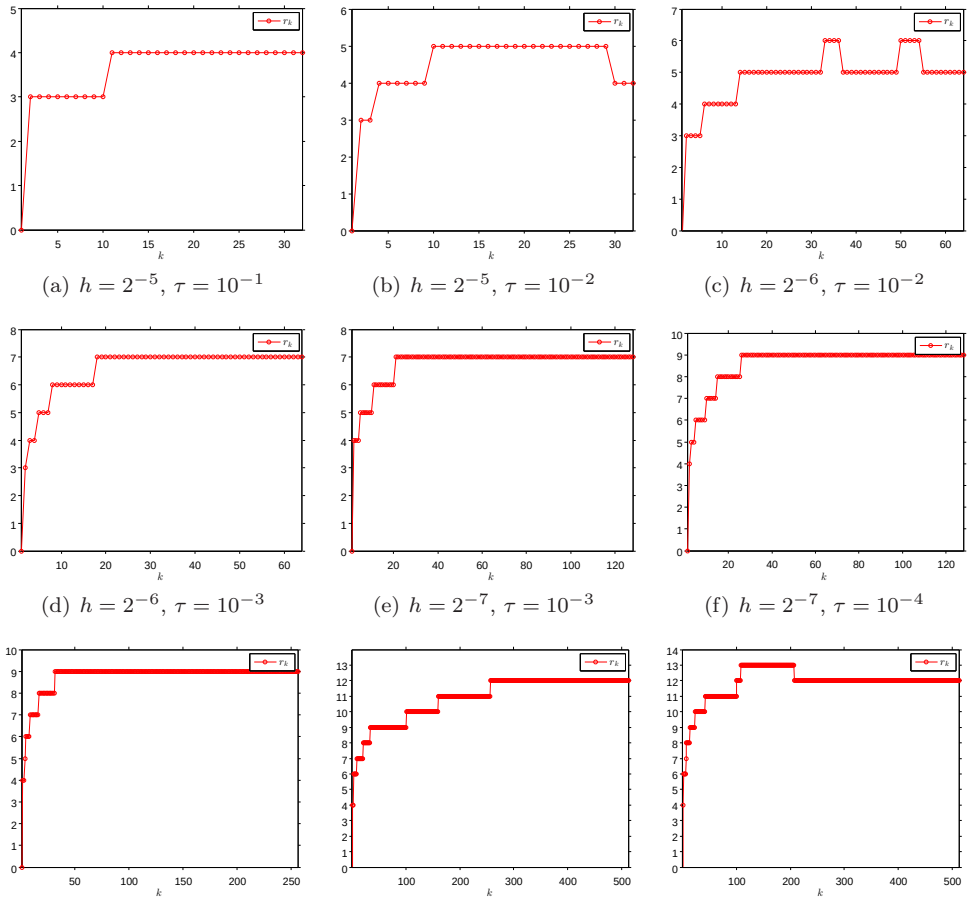
(e) $h = 2^{-7}$, $\tau = 10^{-1}$ (f) $h = 2^{-7}$, $\tau = 10^{-2}$

(g) $h = 2^{-8}$, $\tau = 10^{-2}$ (h) $h = 2^{-8}$, $\tau = 10^{-3}$ (i) $h = 2^{-9}$, $\tau = 10^{-3}$

FIG. 34. *Adaptive semiseparable order for* $\beta = 10^{-2}$

For much smaller $\beta = 10^{-5}$, the saddle-point system is even more ill-conditioned. The smallest singular value for all the principle leading sub-matrices is even more smaller. To solve such an ill-conditioned system, a smaller τ is necessary to get a satisfactory radius of the circle that contains the preconditioned spectrum, compared

with the case for moderate β. This yields a slightly increase of the semiseparable order and the semiseparable orders for all the test cases are still bounded by a small constant, which is illustrated by Figure 35.



(a) $h = 2^{-5}$, $\tau = 10^{-1}$ (b) $h = 2^{-5}$, $\tau = 10^{-2}$ (c) $h = 2^{-6}$, $\tau = 10^{-2}$

(d) $h = 2^{-6}$, $\tau = 10^{-3}$ (e) $h = 2^{-7}$, $\tau = 10^{-3}$ (f) $h = 2^{-7}$, $\tau = 10^{-4}$

(g) $h = 2^{-8}$, $\tau = 10^{-4}$ (h) $h = 2^{-9}$, $\tau = 10^{-4}$ (i) $h = 2^{-9}$, $\tau = 10^{-5}$

FIG. 35. *Adaptive semiseparable order for* $\beta = 10^{-5}$

REMARK 6.2. According to the computational results for different regularization parameter β and mesh size h in Figure 33-Figure 35 and Table 5, we show that by a properly setting of the parameter τ for the MSSS preconditioner, we can compute an efficient preconditioner that gives mesh size and regularization parameter independent convergence. The computational complexity for the MSSS preconditioning technique can be kept linear with the problem sizes.

7. Conclusions. In this manuscript, we made a convergence analysis of the multilevel sequentially semiseparable (MSSS) preconditioners for a wide class of linear systems. This includes unsymmetric systems, symmetric indefinite systems from discretization of scalar PDEs and saddle-point systems. We showed that the spectrum of the preconditioned system is contained in a circle centered at $(1, 0)$ and we gave an analytic bound for the radius. Our analysis shows that the radius of the circle can be made arbitrarily small by properly setting a parameter in the MSSS preconditioner. We also demonstrated how to select the parameter. We validate our analysis

by performing numerical experiments.

The next step of our research is to focus on applying the MSSS preconditioning technique to the wind farm control. This type of application belongs to the in-domain control problems and yields a linear system of the saddle-point type. Standard block preconditioners fail to solve this problem since the Schur complement is difficult or even impossible to approximate. By solving this saddle-point system using our MSSS preconditioning technique, we can obtain the optimal control to maximize the total output power of a wind farm.

Appendix A. Proof of Lemma 5.2. Before the proof, we give the following lemmas and corollaries that are necessary.

LEMMA A.1 ([41]). *Let $A \in \mathbb{C}^{m\times n}$ be partitioned in the form*

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}.$$

Let the singular values of A be $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ and those of A_1 be $\tau_1 \geq \tau_2 \geq \cdots \geq \tau_n$. Then

$$\sigma_i \geq \tau_i, \quad i = 1,\ 2,\ \ldots,\ n.$$

From Lemma A.1, we can also get the inequality between singular values of A and A_2, which is stated by Proposition A.2.

PROPOSITION A.2. *Let the singular values of A_2 in Lemma A.1 be $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_n$. Then*

$$\sigma_i \geq \nu_i, \quad i = 1,\ 2,\ \ldots,\ n.$$

Proof. It is easy to obtain

$$\begin{bmatrix} A_2 \\ A_1 \end{bmatrix} = \begin{bmatrix} 0 & I_p \\ I_q & 0 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} 0 & I_p \\ I_q & 0 \end{bmatrix} A,$$

where I_p and I_q are identity matrices with proper sizes. Let $\bar{A} = \begin{bmatrix} 0 & I_p \\ I_q & 0 \end{bmatrix} A$, then according to Lemma A.1, we have

$$\sigma_i \geq \nu_i, \quad i = 1,\ 2,\ \ldots,\ n.$$

This is because that $\bar{A}$ and A have the same singular values. ∎

According to Lemma A.1 and Proposition A.2, we have the following corollary.

COROLLARY A.3. *If all the factors $\mathcal{C}_i$ are transformed to the form with orthonormal rows, then we have*

$$\|R_i\|_2 \leq 1, \ \textit{and}\ \|Q_i\|_2 \leq 1. \tag{A.1}$$

Proof. According to the procedure to transform $\mathcal{C}_i$ to the form with orthonormal rows introduced in Section 5, at step $i+1$, we perform an SVD that gives

$$\begin{bmatrix} R_i & Q_i \end{bmatrix} = U_i \Sigma_i V_i^T,$$

and let $\begin{bmatrix} R_i & Q_i \end{bmatrix} = V_i^T$. This gives

$$V_i = \begin{bmatrix} R_i^T \\ Q_i^T \end{bmatrix}.$$

According to Lemma A.1 and Proposition A.2, we have

$$\sigma_k(V_i) \geq \sigma_k(R_i^T), \quad \sigma_k(V_i) \geq \sigma_k(Q_i^T).$$

Since $\sigma_k(R_i) = \sigma_k(R_i^T)$, $\sigma_k(Q_i) = \sigma_k(Q_i^T)$ and $\sigma_k(V_i) = 1$, we have

$$\|R_i\|_2 \leq 1, \quad \|Q_i\|_2 \leq 1. \qquad \square$$

With these lemmas and corollaries, we now give the proof of Lemma 5.2 in the following part.

Proof. Since these inequalities in the lemma start from different steps, we first give the proof at step N and then start the proof by induction from step $N-1$.

For step N, perform an SVD on $\mathcal{O}_N$ gives

$$\mathcal{O}_N = P_N = \begin{bmatrix} U_N & \Delta U_N \end{bmatrix} \begin{bmatrix} \Sigma_N & \\ & \Delta\Sigma_N \end{bmatrix} \begin{bmatrix} V_N^T \\ \Delta V_N^T \end{bmatrix},$$

where Σ_N and $\Delta\Sigma_N$ are diagonal matrices with diagonal entries $\sigma_1, \ \sigma_2, \ \cdots, \ \sigma_{\tilde{r}_N}$, and $\sigma_{\tilde{r}_{N+1}}, \ \cdots, \ \sigma_{r_N}$ with

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\tilde{r}_N} > \tau \geq \sigma_{\tilde{r}_{N+1}} \geq \cdots \geq \sigma_{r_N}.$$

Let $\tilde{\mathcal{O}}_N = U_N$ and $\tilde{\mathcal{C}}_N = \Sigma_N V_N^T \mathcal{C}_N$, we have

$$\begin{aligned} \left\| \mathcal{O}_N \mathcal{C}_N - \tilde{O}_N \tilde{\mathcal{C}}_N \right\|_2 &= \left\| \Delta U_N \Delta\Sigma_N \Delta V_N^T \mathcal{C}_N \right\|_2 \\ &= \left\| \Delta U_N \Delta\Sigma_N \Delta V_N^T \right\|_2 \quad (\mathcal{C}_N \text{ has orthonormal rows}) \\ &\leq \tau. \end{aligned}$$

This is exactly the inequality in (5.1) for $i = N$, i.e.,

$$\left\| \tilde{H}_N - \mathcal{H}_N \right\|_2 \leq \tau.$$

After this step, the factor $\tilde{\mathcal{O}}_N$ has orthonormal columns.

According to $\mathcal{C}_N = \begin{bmatrix} R_{N-1}\mathcal{C}_{N-1} & Q_{N-1} \end{bmatrix}$ and $\tilde{\mathcal{C}}_N = \Sigma_N V_N^T \mathcal{C}_N$, we have

$$\tilde{R}_{N-1}^1 = \Sigma_N V_N^T R_{N-1}, \quad \tilde{Q}_{N-1} = \Sigma_N V_N^T Q_{N-1}.$$

Then

$$\begin{aligned} \left\| \tilde{\mathcal{O}}_N \tilde{Q}_{N-1} - \mathcal{O}_N Q_{N-1} \right\|_2 &= \left\| \Delta U_N \Delta\Sigma_N \Delta V_N^T Q_{N-1} \right\|_2 \\ &\leq \left\| \Delta U_N \Delta\Sigma_N \Delta V_N^T \right\|_2 \left\| Q_{N-1} \right\|_2 \\ &\leq \left\| \Delta U_N \Delta\Sigma_N \Delta V_N^T \right\|_2 \quad (\text{Corollary A.3}) \\ &\leq \tau, \end{aligned}$$

which gives the inequality (5.3) for $i = N$.

Now, we start our proof from the step $i = N-1$ by induction. Because of the approximation of $\tilde{\mathcal{O}}_N$, we have

$$\tilde{\mathcal{O}}^1_{N-1} = \begin{bmatrix} P_{N-1} \\ \tilde{\mathcal{O}}_N \tilde{R}^1_{N-1} \end{bmatrix}$$

and we have the following inequality hold

$$\left\| \mathcal{O}_{N-1} - \tilde{\mathcal{O}}^1_{N-1} \right\|_2 \leq \tau. \tag{A.2}$$

This is by the reason of

$$\begin{aligned}
\left\| \mathcal{O}_{N-1} - \tilde{\mathcal{O}}^1_{N-1} \right\|_2 &= \left\| \begin{bmatrix} P_{N-1} \\ \mathcal{O}_N R_{N-1} \end{bmatrix} - \begin{bmatrix} P_{N-1} \\ \tilde{\mathcal{O}}_N \tilde{R}^1_{N-1} \end{bmatrix} \right\|_2 \\
&= \left\| \begin{bmatrix} 0 \\ \mathcal{O}_N R_{N-1} - \tilde{\mathcal{O}}_N \tilde{R}^1_{N-1} \end{bmatrix} \right\|_2 \\
&= \left\| \mathcal{O}_N R_{N-1} - \tilde{\mathcal{O}}_N \tilde{R}^1_{N-1} \right\|_2 \\
&= \left\| \Delta U_N \Delta \Sigma_N \Delta V_N^T R_{N-1} \right\|_2 \\
&\leq \left\| \Delta U_N \Delta \Sigma_N \Delta V_N^T \right\|_2 \left\| R_{N-1} \right\|_2 \\
&\leq \left\| \Delta U_N \Delta \Sigma_N \Delta V_N^T \right\|_2 \quad \text{(Corollary A.3)} \\
&\leq \tau.
\end{aligned}$$

This proves the inequality (5.2) for $i = N-1$.

According to

$$\tilde{\mathcal{O}}^1_{N-1} = \begin{bmatrix} P_{N-1} \\ \tilde{\mathcal{O}}_N \tilde{R}^1_{N-1} \end{bmatrix} = \begin{bmatrix} I & \\ & \tilde{\mathcal{O}}_N \end{bmatrix} \begin{bmatrix} P_{N-1} \\ \tilde{R}^1_{N-1} \end{bmatrix},$$

and $\tilde{\mathcal{O}}_N$ has orthonormal columns, we perform an SVD on $\begin{bmatrix} P_{N-1} \\ \tilde{R}^1_{N-1} \end{bmatrix}$ and get

$$\begin{bmatrix} P_{N-1} \\ \tilde{R}^1_{N-1} \end{bmatrix} = \begin{bmatrix} U_{N-1} & \Delta U_{N-1} \end{bmatrix} \begin{bmatrix} \Sigma_{N-1} & \\ & \Delta \Sigma_{N-1} \end{bmatrix} \begin{bmatrix} V_{N-1}^T \\ \Delta V_{N-1}^T \end{bmatrix}.$$

Let $\tilde{\mathcal{O}}^2_{N-1} = \begin{bmatrix} I & \\ & \tilde{\mathcal{O}}_N \end{bmatrix} U_{N-1}$ and $\tilde{\mathcal{C}}_{N-1} = \Sigma_{N-1} V_{N-1}^T \mathcal{C}_{N-2}$, i.e.,

$$\tilde{\mathcal{C}}_{N-1} = \Sigma_{N-1} V_{N-1}^T \begin{bmatrix} R_{N-2}\mathcal{C}_{N-2} & Q_{N-2} \end{bmatrix},$$

which yields $\tilde{R}^1_{N-2} = \Sigma_{N-1} V_{N-1}^T R_{N-2}$ and $\tilde{Q}_{N-2} = \Sigma_{N-1} V_{N-1}^T Q_{N-2}$. Then, we obtain

$$\begin{aligned}
\left\| \tilde{\mathcal{O}}^2_{N-1} \tilde{\mathcal{C}}_{N-1} - \tilde{\mathcal{O}}^1_{N-1} \mathcal{C}_{N-1} \right\|_2 &= \left\| \Delta U_{N-1} \Delta \Sigma_{N-1} \Delta V_{N-1}^T \mathcal{C}_{N-1} \right\|_2 \\
&= \left\| \Delta U_{N-1} \Delta \Sigma_{N-1} \Delta V_{N-1}^T \right\|_2 \quad (\mathcal{C}_{N-1} \text{ has orthonormal rows}) \\
&\leq \tau.
\end{aligned} \tag{A.3}$$

This in turn gives

$$\begin{aligned}\left\|\mathcal{O}_{N-1}\mathcal{C}_{N-1}-\tilde{\mathcal{O}}^2_{N-1}\tilde{\mathcal{C}}_{N-1}\right\|_2 &\le \left\|\mathcal{O}_{N-1}\mathcal{C}_{N-1}-\tilde{\mathcal{O}}^1_{N-1}\mathcal{C}_{N-1}\right\|_2+\left\|\tilde{\mathcal{O}}^1_{N-1}\mathcal{C}_{N-1}-\tilde{\mathcal{O}}^2_{N-1}\tilde{\mathcal{C}}_{N-1}\right\|_2\\ &=\left\|\mathcal{O}_{N-1}-\tilde{\mathcal{O}}^1_{N-1}\right\|_2+\left\|\tilde{\mathcal{O}}^1_{N-1}\mathcal{C}_{N-1}-\tilde{\mathcal{O}}^2_{N-1}\tilde{\mathcal{C}}_{N-1}\right\|_2\\ &\le 2\tau \quad (\mathcal{C}_{N-1}\text{ has orthonormal rows and (A.2) (A.3)}),\end{aligned}$$

i.e.,

$$\left\|\mathcal{H}_{N-1}-\tilde{\mathcal{H}}_{N-1}\right\|_2\le 2\tau,$$

which proves inequality (5.1) for $i=N-1$.

Additionally,

$$\begin{aligned}\left\|\tilde{\mathcal{O}}^2_{N-1}\tilde{Q}_{N-2}-\tilde{\mathcal{O}}^1_{N-1}Q_{N-2}\right\|_2 &=\left\|\begin{bmatrix} I & \\ & \tilde{\mathcal{O}}_N\end{bmatrix}\Delta U_{N-1}\Delta\Sigma_{N-1}\Delta V^T_{N-1}Q_{N-2}\right\|_2\\ &=\left\|\Delta U_{N-1}\Delta\Sigma_{N-1}\Delta V^T_{N-1}Q_{N-2}\right\|_2 \ (\tilde{\mathcal{O}}_N\text{ has orthonormal columns})\\ &\le\left\|\Delta U_{N-1}\Delta\Sigma_{N-1}\Delta V^T_{N-1}\right\|_2, \quad \text{(Corollary A.3)}\\ &\le\tau.\end{aligned}$$

And,

$$\begin{aligned}\left\|\tilde{\mathcal{O}}^1_{N-1}Q_{N-2}-\mathcal{O}_{N-1}Q_{N-2}\right\|_2 &\le\left\|\tilde{\mathcal{O}}^1_{N-1}-\mathcal{O}_{N-1}\right\|_2\|Q_{N-2}\|_2\\ &\le\left\|\tilde{\mathcal{O}}^1_{N-1}-\mathcal{O}_{N-1}\right\|_2 \quad \text{(Corollary A.3)}\\ &\le\tau. \quad \text{(Equation (A.2))}\end{aligned}$$

This in turn yields

(A.4)
$$\begin{aligned}\left\|\tilde{\mathcal{O}}^2_{N-1}\tilde{Q}_{N-2}-\mathcal{O}_{N-1}Q_{N-2}\right\|_2 &\le\left\|\tilde{\mathcal{O}}^2_{N-1}\tilde{Q}_{N-2}-\tilde{\mathcal{O}}^1_{N-1}Q_{N-2}\right\|_2+\left\|\tilde{\mathcal{O}}^1_{N-1}Q_{N-2}-\mathcal{O}_{N-1}Q_{N-2}\right\|_2\\ &\le 2\tau,\end{aligned}$$

which is exactly inequality (5.3) for $i=N-1$.

Till now, we have proven that all the inequalities (5.1) (5.2) (5.3) hold for $i=N-1$. Next, we suppose that at step $(k+1)$, $2\le k\le N-2$, the following inequalities hold,

$$\begin{aligned}\left\|\mathcal{O}^1_{k+1}-\tilde{\mathcal{O}}^1_{k+1}\right\|_2&\le(N-k-1)\tau,\\ \left\|\tilde{\mathcal{O}}^2_{k+1}\tilde{Q}_k-\mathcal{O}_{k+1}Q_k\right\|_2&\le(N-k)\tau.\end{aligned}$$

Therefore, at step k, we have

$$\begin{aligned}\left\|\tilde{\mathcal{O}}^1_k-\mathcal{O}_k\right\|_2 &=\left\|\begin{bmatrix}P_k\\ \tilde{\mathcal{O}}^2_{k+1}\tilde{R}^1_k\end{bmatrix}-\begin{bmatrix}P_k\\ \mathcal{O}_{k+1}R_k\end{bmatrix}\right\|_2\\ &=\left\|\tilde{\mathcal{O}}^2_{k+1}\tilde{R}^1_k-\mathcal{O}_{k+1}R_k\right\|_2\\ &\le\left\|\tilde{\mathcal{O}}^2_{k+1}\tilde{R}^1_k-\tilde{\mathcal{O}}^1_{k+1}R_k\right\|_2+\left\|\tilde{\mathcal{O}}^1_{k+1}R_k-\mathcal{O}_{k+1}R_k\right\|_2.\end{aligned}$$

And it is easy to obtain

$$\begin{aligned}\left\|\tilde{\mathcal{O}}_{k+1}^2\tilde{R}_k^1-\tilde{\mathcal{O}}_{k+1}^1R_k\right\|_2 &= \left\|\begin{bmatrix} I & \\ & \tilde{\mathcal{O}}_{k+1}^2\end{bmatrix}\Delta U_{k+1}\Delta\Sigma_{k+1}\Delta V_{k+1}^TR_k\right\|_2\\ &= \left\|\Delta U_{k+1}\Delta\Sigma_{k+1}\Delta V_{k+1}^TR_k\right\|_2 \quad (\tilde{\mathcal{O}}_{k+1}^2 \text{ has orthonormal columns})\\ &\le \left\|\Delta U_{k+1}\Delta\Sigma_{k+1}\Delta V_{k+1}^T\right\|_2 \quad (\text{Corollary A.3})\\ &\le \tau.\end{aligned}$$

Besides

$$\begin{aligned}\left\|\tilde{\mathcal{O}}_{k+1}^1R_k-\mathcal{O}_{k+1}R_k\right\|_2 &\le \left\|\tilde{\mathcal{O}}_{k+1}^1-\mathcal{O}_{k+1}\right\|_2\|R_k\|_2\\ &\le \left\|\tilde{\mathcal{O}}_{k+1}^1-\mathcal{O}_{k+1}\right\|_2 \quad (\text{Corollary A.3})\\ &\le (N-k-1)\tau.\end{aligned}$$

This gives

$$\left\|\tilde{\mathcal{O}}_k^1-\mathcal{O}_k\right\|_2 \le \tau+(N-k-1)\tau=(N-k)\tau,$$

which is exactly the inequality (5.2) for step k.

Next, we start to approximate $\tilde{\mathcal{O}}_k^1$. Since

$$\tilde{\mathcal{O}}_k^1=\begin{bmatrix}P_k\\ \tilde{\mathcal{O}}_{k+1}^2\tilde{R}_k^1\end{bmatrix}=\begin{bmatrix} I & \\ & \tilde{\mathcal{O}}_{k+1}^2\end{bmatrix}\begin{bmatrix}P_k\\ \tilde{R}_k^1\end{bmatrix},$$

and $\tilde{\mathcal{O}}_{k+1}^2$ has orthonormal columns, we first perform an SVD on $\begin{bmatrix}P_k\\ \tilde{R}_k^1\end{bmatrix}$ that gives

$$\begin{bmatrix}P_k\\ \tilde{R}_k^1\end{bmatrix}=\begin{bmatrix}U_k & \Delta U_k\end{bmatrix}\begin{bmatrix}\Sigma_k & \\ & \Delta\Sigma_k\end{bmatrix}\begin{bmatrix}V_k^T\\ \Delta V_k^T\end{bmatrix}.$$

Let

$$\begin{bmatrix}\tilde{P}_k\\ \tilde{R}_k\end{bmatrix}=U_k, \text{ and } \tilde{\mathcal{C}}_k=\Sigma_kV_k^T\mathcal{C}_k.$$

This yields,

$$\begin{aligned}\left\|\tilde{\mathcal{O}}_k^2\tilde{\mathcal{C}}_k-\tilde{\mathcal{O}}_k^1\mathcal{C}_k\right\|_2 &= \left\|\Delta U_k\Delta\Sigma_k\Delta V_k^T\mathcal{C}_k\right\|_2\\ &= \left\|\Delta U_k\Delta\Sigma_k\Delta V_k^T\right\|_2 \quad (\mathcal{C}_k \text{ has orthonormal rows})\\ &\le \tau.\end{aligned}$$

Therefore,

$$\begin{aligned}\left\|\tilde{\mathcal{O}}_k^2\tilde{\mathcal{C}}_k-\mathcal{O}_k\mathcal{C}_k\right\|_2 &\le \left\|\tilde{\mathcal{O}}_k^2\tilde{\mathcal{C}}_k-\tilde{\mathcal{O}}_k^1\mathcal{C}_k\right\|_2+\left\|\tilde{\mathcal{O}}_k^1\mathcal{C}_k-\mathcal{O}_k\mathcal{C}_k\right\|_2\\ &\le \tau+\left\|\tilde{\mathcal{O}}_k^1\mathcal{C}_k-\mathcal{O}_k\mathcal{C}_k\right\|_2\\ &= \tau+\left\|\tilde{\mathcal{O}}_k^1-\mathcal{O}_k\right\|_2 \quad (\mathcal{C}_k \text{ has orthonormal rows})\\ &\le (N-k+1)\tau,\end{aligned}$$

i.e.,

$$\left\|\tilde{\mathcal{H}}_k - \mathcal{H}_k\right\|_2 \leq (N-k+1)\tau,$$

which proves inequality (5.1) for step k.

Besides,

$$\begin{aligned}\left\|\tilde{\mathcal{O}}_k^2\tilde{Q}_{k-1} - \mathcal{O}_k Q_{k-1}\right\|_2 &\leq \left\|\tilde{\mathcal{O}}_k^2\tilde{Q}_{k-1} - \tilde{\mathcal{O}}_k^1 Q_{k-1}\right\|_2 + \left\|\tilde{\mathcal{O}}_k^1 Q_{k-1} - \mathcal{O}_k Q_{k-1}\right\|_2 \\ &\leq \left\|\tilde{\mathcal{O}}_k^2\tilde{Q}_{k-1} - \tilde{\mathcal{O}}_k^1 Q_{k-1}\right\|_2 + \left\|\tilde{\mathcal{O}}_k^1 - \mathcal{O}_k\right\|_2. \quad \text{(Corollary A.3)}\end{aligned}$$

It is easy to obtain that

$$\begin{aligned}\left\|\tilde{\mathcal{O}}_k^2\tilde{Q}_{k-1} - \tilde{\mathcal{O}}_k^1 Q_{k-1}\right\|_2 &= \left\|\begin{bmatrix} I & \\ & \tilde{\mathcal{O}}_{k+1}^2\end{bmatrix}\Delta U_k \Delta\Sigma_k \Delta V_k^T Q_{k-1}\right\|_2 \\ &= \left\|\Delta U_k \Delta\Sigma_k \Delta V_k^T Q_{k-1}\right\|_2 \quad (\tilde{\mathcal{O}}_{k+1}^2 \text{ has orthonormal columns}) \\ &\leq \left\|\Delta U_k \Delta\Sigma_k \Delta V_k^T\right\|_2 \left\|Q_{k-1}\right\|_2 \\ &\leq \left\|\Delta U_k \Delta\Sigma_k \Delta V_k^T\right\|_2 \quad \text{(Corollary A.3)} \\ &\leq \tau.\end{aligned}$$

Therefore,

$$\left\|\tilde{\mathcal{O}}_k^2\tilde{Q}_{k-1} - \mathcal{O}_k Q_{k-1}\right\|_2 \leq \tau + (N-k)\tau = (N-k+1)\tau.$$

This gives the proof of inequality (5.3) for step k. ∎

REFERENCES

[1] Hakan Bačcl, Joseph E. Pasciak, and Kostyantyn Y. Sirenko, *A convergence analysis for a sweeping preconditioner for block tridiagonal systems of linear equations*, Numerical Linear Algebra with Applications, 22 (2015), pp. 371–392.

[2] Mario Bebendorf, *Why finite element discretizations can be factored by triangular hierarchical matrices*, SIAM Journal on Numerical Analysis, 45 (2007), pp. 1472–1494.

[3] Michele Benzi, Gene H. Golub, and Jorg Liesen, *Numerical solution of saddle point problems*, Acta Numerica, 14 (2005), pp. 1–137.

[4] Dennis S. Bernstein, *Matrix mathematics: theory, facts, and formulas*, Princeton University Press, 2009.

[5] Steffen Börm, *$\mathcal{H}^2$-matrices–multilevel methods for the approximation of integral operators*, Computing and Visualization in Science, 7 (2004), pp. 173–181.

[6] Steffen Börm and Sabine Le Borne, *$\mathcal{H}$-LU factorization in preconditioners for augmented Lagrangian and grad-div stabilized saddle point systems*, International Journal for Numerical Methods in Fluids, 68 (2012), pp. 83–98.

[7] Shivkumar Chandrasekaran, Patrick Dewilde, Ming Gu, William Lyons, and Timothy Pals, *A fast solver for HSS representations via sparse matrices*, SIAM Journal on Matrix Analysis and Applications, 29 (2006), pp. 67–81.

[8] Shivkumar Chandrasekaran, Patrick Dewilde, Ming Gu, Timothy P. Pals, Xiaorui Sun, Alle-Jan van der Veen, and Daniel White, *Some fast algorithms for sequentially semiseparable representations*, SIAM Journal on Matrix Analysis and Applications, 27 (2005), pp. 341–364.

[9] Shivkumar Chandrasekaran, Patrick Dewilde, Ming Gu, and Naveen Somasunderam, *On the numerical rank of the off-diagonal blocks of Schur complements of discretized elliptic PDEs*, SIAM Journal on Matrix Analysis and Applications, 31 (2010), pp. 2261–2290.

[10] Paul Concus, Gene Golub, and Gérard Meurant, *Block preconditioning for the conjugate gradient method*, SIAM Journal on Scientific and Statistical Computing, 6 (1985), pp. 220–252.

[11] Patrick Dewilde, Haiyan Jiao, and Shivkuma Chandrasekaran, *Model reduction in symbolically semi-separable systems with application to preconditioners for 3D sparse systems of equations*, in Characteristic Functions, Scattering Functions and Transfer Functions, vol. 197 of Operator Theory: Advances and Applications, Birkhäser Basel, 2010, pp. 99–132.

[12] Patrick Dewilde and Alle-Jan van der Veen, *Time-varying systems and computations*, Kluwer Academic Publishers, Boston, 1998.

[13] Yuli Eidelman and Israel Gohberg, *On a new class of structured matrices*, Integral Equations and Operator Theory, 34 (1999), pp. 293–324.

[14] ———, *On generators of quasiseparable finite block matrices*, Calcolo, 42 (2005), pp. 187–214.

[15] Howard C. Elman, David J. Silvester, and Andrew J. Wathen, *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics*, Oxford University Press, New York, 2005.

[16] Björn Engquist and Lexing Ying, *Sweeping preconditioner for the Helmholtz equation: hierarchical matrix representation*, Communications on pure and applied mathematics, 64 (2011), pp. 697–735.

[17] Yogi A. Erlangga, Cornelis W. Oosterlee, and Cornelis Vuik, *A novel multigrid based preconditioner for heterogeneous Helmholtz problems*, SIAM Journal on Scientific Computing, 27 (2006), pp. 1471–1492.

[18] Gene H. Golub and Charles F. Van Loan, *Matrix computations*, Johns Hopkins University Press, Baltimore, 1996.

[19] Ming Gu, Xiaoye S. Li, and Panayot S. Vassilevski, *Direction-preserving and Schur-monotonic semiseparable approximations of symmetric positive definite matrices*, SIAM Journal on Matrix Analysis and Applications, 31 (2010), pp. 2650–2664.

[20] Wolfgang Hackbusch, *A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices*, Computing, 62 (1999), pp. 89–108.

[21] Wolfgang Hackbusch and Steffen Börm, *Data-sparse approximation by adaptive $\mathcal{H}^2$-matrices*, Computing, 69 (2002), pp. 1–35.

[22] Sabine Le Borne and Lars Grasedyck, *$\mathcal{H}$-matrix preconditioners in convection-dominated problems*, SIAM Journal on Matrix Analysis and Applications, 27 (2006), pp. 1172–1183.

[23] Gérard Meurant, *A review on the inverse of symmetric tridiagonal and block tridiagonal matrices*, SIAM Journal on Matrix Analysis and Applications, 13 (1992), pp. 707–728.

[24] Artem Napov, *Conditioning analysis of incomplete Cholesky factorizations with orthogonal dropping*, SIAM Journal on Matrix Analysis and Applications, 34 (2013), pp. 1148–1173.

[25] Cornelis W. Oosterlee, Cornelis Vuik, W.A. Mulder, and R.-E. Plessix, *Shifted-Laplacian preconditioners for heterogeneous Helmholtz problems*, in Advanced Computational Methods in Science and Engineering, Barry Koren and Kees Vuik, eds., vol. 71 of Lecture Notes in Computational Science and Engineering, Springer-Verlag Berlin Heidelberg, 2010, pp. 21–46.

[26] Chris Paige and Michael Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM Journal on Numerical Analysis, 12 (1975), pp. 617–629.

[27] John W. Pearson, *On the development of parameter-robust preconditioners and commutator arguments for solving stokes control problems*, Electronic Transactions on Numerical Analysis, 44 (2015), pp. 53–72.

[28] Yue Qiu, Martin B. van Gijzen, Jan-Willem van Wingerden, and Michel Verhaegen, *A class of efficient preconditioners with multilevel sequentially semiseparable matrix structure*, AIP Conference Proceedings, 1558 (2013), pp. 2253–2256.

[29] Yue Qiu, Martin B. van Gijzen, Jan-Willem van Wingerden, Michel Verhaegen, and Cornelis Vuik, *Efficient preconditioners for PDE-constrained optimization problems with a multilevel sequentially semiseparable matrix structure*, Tech. Report 13-04, Delft Institution of Applied Mathematics, Delft University of Technology, 2013. Available at `http://ta.twi.tudelft.nl/nw/users/yueqiu/publications.html`.

[30] ———, *Convergence analysis of the multilevel sequentially semiseparable preconditioners*, Tech. Report 15-01, Delft Institution of Applied Mathematics, Delft University of Technology, 2015. Available at `http://ta.twi.tudelft.nl/nw/users/yueqiu/publications.html`.

[31] ———, *Efficient preconditioners for PDE-constrained optimization problems with a multilevel sequentially semiseparable matrix structure*, Electronic Transactions on Numerical Analysis, (2015). Available at `http://ta.twi.tudelft.nl/nw/users/yueqiu/publications.html`.

[32] ———, *Evaluation of multilevel sequentially semiseparable preconditioners on CFD benchmark problems using incompressible flow and iterative solver software*, Mathematical Methods in the Applied Sciences, 38 (2015), pp. n/a–n/a.

[33] Tyrone Rees, *Preconditioning Iterative Methods for PDE-Constrained Optimization*, PhD thesis, University of Oxford, 2010.

[34] Tyrone Rees and Andrew J. Wathen, *Preconditioning iterative methods for the optimal control of the stokes equations*, SIAM Journal on Scientific Computing, 33 (2011), pp. 2903–2926.

[35] Justin K. Rice, *Efficient Algorithms for Distributed Control: a Structured Matrix Approach*, PhD thesis, Delft University of Technology, 2010.

[36] Yousef Saad, *Iterative methods for sparse linear systems*, Society for Industrial and Applied Mathematics, Philadelphia, 2003.

[37] Abdul H. Sheikh, Domenico Lahaye, and Cornelis Vuik, *On the convergence of shifted Laplace preconditioner combined with multilevel deflation*, Numerical Linear Algebra with Applications, 20 (2013), pp. 645–662.

[38] David J. Silvester, Howard C. Elman, and Alison Ramage, *Incompressible Flow and Iterative Solver Software (IFISS) version 3.2*, May 2012. `http://www.manchester.ac.uk/ifiss/`.

[39] Peter Sonneveld and Martin B. van Gijzen, *IDR(s): A family of simple and fast algorithms for solving large nonsymmetric systems of linear equations*, SIAM Journal on Scientific Computing, 31 (2008), pp. 1035–1062.

[40] Gilbert W. Stewart, *Perturbation theory for the singular value decomposition*, tech. report, College Park, MD, USA, 1990.

[41] Gilbert W. Stewart, Jiguang Sun, and Harcourt B. Jovanovich, *Matrix perturbation theory*, Academic Press, New York, 1990.

[42] Martin Stoll and Tobias Breiten, *A low-rank in time approach to PDE-constrained optimization*, SIAM Journal on Scientific Computing, 37 (2015), pp. B1–B29.

[43] Martin B. van Gijzen, Yogi A. Erlangga, and Cornelis Vuik, *Spectral analysis of the discrete Helmholtz operator preconditioned with a shifted Laplacian*, SIAM Journal on Scientific Computing, 29 (2007), pp. 1942–1958.

[44] Martin B. van Gijzen and Peter Sonneveld, *Algorithm 913: An elegant IDR(s) variant that efficiently exploits biorthogonality properties*, ACM Transactions on Mathematical Software, 38 (2011), pp. 5:1–5:19.

[45] Raf Vandebril, Marc Van Barel, and Nicola Mastronardi, *Matrix computations and semiseparable matrices: linear systems*, Johns Hopkins University Press, Baltimore, 2007.

[46] Andy Wathen and David Silvester, *Fast iterative solution of stabilised Stokes systems. Part I: Using simple diagonal preconditioners*, SIAM Journal on Numerical Analysis, 30 (1993), pp. 630–649.

[47] Jianlin Xia, *A robust inner-outer hierarchically semi-separable preconditioner*, Numerical Linear Algebra with Applications, 19 (2012), pp. 992–1016.

[48] Jianlin Xia, Shivkumar Chandrasekaran, Ming Gu, and Xiaoye Li, *Superfast multifrontal method for large structured linear systems of equations*, SIAM Journal on Matrix Analysis and Applications, 31 (2010), pp. 1382–1411.