

A BLOCK TOEPLITZ PRECONDITIONER FOR ALL-AT-ONCE SYSTEMS FROM LINEAR WAVE EQUATIONS*

SEAN HON[†] AND STEFANO SERRA-CAPIZZANO[‡]

Abstract. In this work, we propose a novel parallel-in-time preconditioner for an all-at-once system, arising from the numerical solution of linear wave equations. Namely, our main result concerns a block tridiagonal Toeplitz preconditioner that can be diagonalized via fast sine transforms, whose effectiveness is theoretically shown for the nonsymmetric block Toeplitz system resulting from discretizing the concerned wave equation. Our approach is to first transform the original linear system into a symmetric one and subsequently develop the desired preconditioning strategy based on the spectral symbol of the modified matrix. Various Krylov subspace methods are considered. That is, we show that the minimal polynomial of the preconditioned matrix is of low degree, which leads to fast convergence when the generalized minimal residual method is used. To fully utilize the symmetry of the modified matrix, we additionally construct an absolute-value preconditioner which is symmetric positive definite. Then, we show that the eigenvalues of the preconditioned matrix are clustered around ± 1 , which gives a convergence guarantee when the minimal residual method is employed. Numerical examples are given to support the effectiveness of our preconditioner. Our block Toeplitz preconditioner provides an alternative to the existing block circulant preconditioner proposed by McDonald, Pestana, and Wathen in [SIAM J. Sci. Comput., 40 (2018), pp. A1012–A1033], advancing the symmetrization preconditioning theory that originated from the same work.

Key words. fast sine transforms, wave equations, Krylov subspace methods, all-at-once discretization, parallel-in-time, block circulant preconditioners

AMS subject classifications. 15B05, 65F08, 65F10, 65M22

1. Introduction. Since [33] there has been an increasing number of works on developing effective preconditioners for solving the all-at-once linear systems stemming from solving evolutionary partial differential equations (PDEs) over the past few years, including [18, 19, 25] for heat equations, [6, 28, 39] for wave equations, and various other differential equations [26, 32, 42]. Instead of solving PDEs in a sequential fashion, these parallel-in-time (PinT) methods solve all unknowns simultaneously by constructing a monolithic linear system which is composed of smaller systems at each time level. These preconditioning solvers belong to the diagonalization-based all-at-once methods [13, 30]; see [14] for an overview of the methods. Other related PinT methods include space-time multigrid [15, 21], multigrid reduction in time [8, 10], and parareal methods [16, 27]. For a survey on the development of PinT solvers, we refer the reader to [12] and the references therein.

In this work, we are interested in developing a preconditioning PinT method for solving the linear wave equation

$$(1.1) \quad \begin{cases} u_{tt}(x, t) = \Delta u(x, t) + f(x, t), & (x, t) \in \Omega \times (0, T], \\ u = 0, & (x, t) \in \partial\Omega \times (0, T], \\ u(x, 0) = \psi_0, \quad u_t(x, 0) = \psi_1, & x \in \Omega. \end{cases}$$

There are a number of closely related all-at-once preconditioning methods proposed for (1.1), including [6, 18, 28, 39], in which the generalized minimal residual method (GMRES) is used and the proposed preconditioners are constructed based on circulant-type matrices. As will be

*Received February 4, 2022. Accepted December 12, 2022. Published online on February 13, 2023. Recommended by Alessandro Buccini.

[†]Corresponding Author. Department of Mathematics, Hong Kong Baptist University, Hong Kong SAR (seanyshon@hkbu.edu.hk).

[‡]Department of Science and High Technology, University of Insubria, Italy (s.serracapizzano@uninsubria.it).

A natural choice of preconditioners for (1.2) is the existing Strang-type block circulant preconditioner proposed in [6, 18, 33] denoted by

$$(1.4) \quad \mathcal{S} = \begin{bmatrix} L_m & & & L_m & -2I_m \\ -2I_m & L_m & & & L_m \\ L_m & -2I_m & L_m & & \\ & \ddots & \ddots & \ddots & \\ & & L_m & -2I_m & L_m \end{bmatrix}.$$

Due to its circulant character, the computation of \mathcal{S}^{-1} using fast Fourier transforms can be parallelized over n difference processors, which makes this approach advantageous in a high performance computing (HPC) setting. However, even though such a block circulant preconditioner was observed effective for well-conditioned block Toeplitz systems, its performance for the ill-conditioned ones is often unsatisfactory as discussed in [20]; see also [7, 38] for theoretical and computational approaches among various matrix-algebras and preconditioning approaches. Thus, this work is motivated to develop a noncirculant preconditioner that is based on the discrete sine transform.

An alternative to directly solving (1.2), we consider the equivalent transformed linear system $\mathcal{Y}\mathcal{T}\mathbf{u} = \mathcal{Y}\mathbf{f}$, where

$$(1.5) \quad \mathcal{Y}\mathcal{T} = \begin{bmatrix} & & L_m & -2I_m & L_m \\ & \ddots & \ddots & \ddots & \\ L_m & -2I_m & L_m & & \\ -2I_m & L_m & & & \\ L_m & & & & \end{bmatrix},$$

and $\mathcal{Y} = Y_n \otimes I_m$ with $Y_n \in \mathbb{R}^{n \times n}$ being the anti-identity matrix (i.e., $[Y_n]_{j,k} = 1$ if and only if $j + k = n + 1$ and $[Y_n]_{j,k} = 0$ otherwise). Clearly, the modified matrix $\mathcal{Y}\mathcal{T}$ is now symmetric.

As will be explained in Section 2, the eigenvalues of $\mathcal{Y}\mathcal{T}$ can be precisely determined by the matrix-valued function $|g|$ up to a \pm sign, where $g = L_m - 2I_m e^{ix} + L_m e^{2ix}$. To match such a spectrum, we propose for $\mathcal{Y}\mathcal{T}$ in this work an SPD preconditioner whose eigenvalues are also related to $|g|$.

Our main result is the following tridiagonal block Toeplitz preconditioner for $\mathcal{Y}\mathcal{T}$:

$$\begin{bmatrix} 2I_m & -L_m & & & \\ -L_m & 2I_m & -L_m & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -L_m \\ & & & -L_m & 2I_m \end{bmatrix} =: \text{BlockToeplitz}(-L_m, 2I_m, -L_m) =: \mathcal{P},$$

which is diagonalized by the discrete sine transform associated to the standard discrete Laplacian. A corresponding preconditioner for the symmetrized system in (1.3) would be $\text{BlockToeplitz}(-I_m, -\tilde{L}_m, -I_m)$. Similar to the block-circulant preconditioner \mathcal{S} , the computation of \mathcal{P}^{-1} using DSTs can also be parallelized over n processors, which is suitable in an HPC system. We refer to [2] for an efficient implementation and [7] for a detailed analysis accounting for the superior performance of sine transform-based preconditioners (also known as τ -preconditioners) over the circulant ones for symmetric Toeplitz systems.

Combined with our proposed preconditioner \mathcal{P} , we will consider various Krylov subspace methods, including GMRES, the minimal residual method (MINRES), and the conjugate

gradient method for the normal equations (CGNE). For GMRES, we will show that the preconditioned matrix $\mathcal{P}^{-1}\mathcal{Y}\mathcal{T}$ has a low-degree minimal polynomial, which implies fast convergence. When CGNE is applied to the normal equation system, we also show similar results for an associated minimal polynomial, which ensures fast convergence.

To fully exploit the symmetry of $\mathcal{Y}\mathcal{T}$, we in addition apply MINRES combined with \mathcal{P} . However, since such a PinT preconditioner \mathcal{P} is not in itself readily SPD, one can construct its so-called “absolute-value” counterpart $|\mathcal{P}| = \sqrt{\mathcal{P}^2}$, which is not only symmetric but also positive definite. We will then prove that a convergence guarantee can be obtained by showing that the preconditioned matrix sequence $\{|\mathcal{P}|^{-1}\mathcal{Y}\mathcal{T}\}_n$ has clustered eigenvalues around ± 1 .

Recently, a τ -preconditioner $\widehat{\mathcal{P}}$ was proposed in our earlier work [19] which was particularly used for heat equations. While also being a τ -preconditioner, our proposed \mathcal{P} in this work is different in several aspects: First, \mathcal{P} is specifically designed for the all-at-once system (1.2) arising from solving the concerned wave equation, while $\widehat{\mathcal{P}}$ is more general, and it can be used for other equations. Second, \mathcal{P} can be used with GMRES while $\widehat{\mathcal{P}}$ is designed mainly for MINRES. Third, unlike $\widehat{\mathcal{P}}$, the construction of \mathcal{P} does not involve the use of the matrix square root. For more about this line of MINRES-approach for heat equations, we refer to [19, 32, 33].

The current paper is organized as follows. In Section 2, we review some preliminary results on block Toeplitz matrices. The main results on our proposed preconditioning methods and their convergence properties are provided in Section 3. Numerical examples are given in Section 4 to support the proposed preconditioners.

2. Preliminaries on block Toeplitz matrices. In this section, we provide some useful background knowledge regarding block Toeplitz matrices.

We let $L^1([-\pi, \pi], \mathbb{C}^{m \times m})$ be the Banach space of all matrix-valued functions that are Lebesgue integrable over $[-\pi, \pi]$. The L^1 -norm induced by the trace norm over $\mathbb{C}^{m \times m}$ is

$$\|f\|_{L^1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \|f(x)\|_{\text{tr}} dx < \infty,$$

where $\|A_n\|_{\text{tr}} := \sum_{j=1}^n \sigma_j(A_n)$ denotes the trace norm of $A_n \in \mathbb{C}^{n \times n}$. The block Toeplitz matrix generated by $f \in L^1([-\pi, \pi], \mathbb{C}^{m \times m})$ is denoted by $T_{(n,m)}[f]$, namely

$$T_{(n,m)} = \begin{bmatrix} A_{(0)} & A_{(-1)} & \cdots & A_{(-n+1)} \\ A_{(1)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_{(-1)} \\ A_{(n-1)} & \cdots & A_{(1)} & A_{(0)} \end{bmatrix} \in \mathbb{C}^{mn \times mn},$$

where the Fourier coefficients of f are

$$A_{(k)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx \in \mathbb{C}^{m \times m}, \quad k = 0, \pm 1, \pm 2, \dots$$

The function f is called the *generating function* of $T_{(n,m)}[f]$. If f is Hermitian, $T_{(n,m)}[f]$ is Hermitian by [36, Lemma 2.1]; in that case f is also the *spectral symbol* of the sequence $\{T_{(n,m)}[f]\}_n$ in the sense of Theorem 2.1, that is, $\{T_{(n,m)}[f]\}_n \sim_{\lambda} f$.

For thorough discussions on the related properties of block Toeplitz matrices, we refer readers to [4, 5, 17, 22, 34] and the references therein. Before discussing the asymptotic spectral distribution of $\mathcal{Y}T_{(n,m)}[f]$ associated with f , it is crucial to our preconditioning theory development that we introduce the following notation:

Given $D \subset \mathbb{R}^k$ with Lebesgue measure $0 < \mu_k(D) < \infty$, we define \tilde{D}_p as $D \cup D_p$, where $p \in \mathbb{R}^k$ and $D_p = p + D$, with the constraint that D and D_p have non-intersecting interior part, i.e., $D^\circ \cap D_p^\circ = \emptyset$. In this way, we have $\mu_k(\tilde{D}_p) = 2\mu_k(D)$. Given any g defined over D , we define ψ_g over \tilde{D}_p in the following fashion:

$$(2.1) \quad \psi_g(x) = \begin{cases} g(x), & x \in D, \\ -g(x-p), & x \in D_p, x \notin D. \end{cases}$$

THEOREM 2.1 ([11, Theorem 3.4], [40]). *Suppose that $f \in L^1([-\pi, \pi], \mathbb{C}^{m \times m})$ with Hermitian Fourier coefficients. Let $T_{(n,m)}[f] \in \mathbb{C}^{mn \times mn}$ be the block Toeplitz matrix generated by f , and let $\mathcal{Y} = Y_n \otimes I_m \in \mathbb{R}^{mn \times mn}$. Then,*

$$(2.2) \quad \{T_{(n,m)}[f]\}_n, \{\mathcal{Y}T_{(n,m)}[f]\}_n \sim_\sigma f,$$

$$(2.3) \quad \{\mathcal{Y}T_{(n,m)}[f]\}_n \sim_\lambda \psi_{|f|}, \quad |f| = (ff^*)^{1/2},$$

over the domain \tilde{D} with $D = [0, 2\pi]$ and $p = -2\pi$, where $\psi_{|f|}$ is defined in (2.1). More precisely, we have $\sigma_j(\mathcal{Y}T_{(n,m)}[f]) = \sigma_j(T_{(n,m)}[f])$, $j = 1, 2, \dots, 2mn$, and

$$(2.4) \quad \lim_{n \rightarrow \infty} \frac{1}{mn} \sum_{j=1}^{mn} F(\sigma_j(\mathcal{Y}T_{(n,m)}[f])) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{m} \sum_{j=1}^m F(\sigma_j(f(x))) dx,$$

$$(2.5) \quad \lim_{n \rightarrow \infty} \frac{1}{mn} \sum_{j=1}^{mn} F(\lambda_j(\mathcal{Y}T_{(n,m)}[f])) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{m} \sum_{j=1}^m F(\lambda_j(\psi_{|f|}(x))) dx,$$

where $\lambda_j(\psi_{|f|}(x))$, $j = 1, 2, \dots, m$, are the eigenvalue functions of $\psi_{|f|}$, $\sigma_j(f)$ are the singular values of f , $j = 1, 2, \dots, m$, and (2.4) is equivalent to (2.2), since $\sigma_j(\mathcal{Y}T_{(n,m)}[f]) = \sigma_j(T_{(n,m)}[f])$, $j = 1, 2, \dots, 2mn$, while (2.5) is equivalent to (2.3).

For the symmetrized matrix $\mathcal{Y}\mathcal{T}$ given in (1.5), we can see by Theorem 2.1 and also a similar result in [31] that its eigenvalues are distributed as $\pm|g|$, where

$$(2.6) \quad g(x) = L_m - 2I_m e^{ix} + L_m e^{2ix},$$

according to the definition in (2.1). Hence, $\mathcal{Y}\mathcal{T}$ is symmetric indefinite when n is sufficiently large, which justifies the use of MINRES.

3. Main results. In this section, we provide the main results on our proposed PinT preconditioner \mathcal{P} defined by (1.6). The design of such a preconditioner is based on the following matrix-valued spectral symbol

$$h = 2(I_m - L_m \cos x) = 2I_m - L_m e^{-ix} - L_m e^{ix},$$

where $|h| = |g|$ with g given in (2.6). Knowing $\mathcal{P} = T_{(n,m)}[h]$, we can readily see that it is a good preconditioner for $\mathcal{Y}\mathcal{T}$ since they share the same absolute-value spectral distribution.

In what follows, we will discuss three Krylov subspace solvers, including GMRES, MINRES, and CGNE, combined with our proposed preconditioner. Related issues such as implementations and convergence analysis for each method are critically discussed.

3.1. GMRES for $\mathcal{P}^{-1}\mathcal{Y}\mathcal{T}$. In this section, we first consider a GMRES solver using our proposed preconditioner \mathcal{P} . We will discuss the implementations of \mathcal{P} and show that the preconditioned matrix $\mathcal{P}^{-1}\mathcal{Y}\mathcal{T}$ has a minimal polynomial of degree independent of the time step n .

3.1.1. Implementations. We begin by discussing the computation of $\mathcal{Y}\mathcal{T}\mathbf{v}$ for any given vector \mathbf{v} . Since \mathcal{T} is a sparse matrix (due to K_m being sparse), computing the matrix-vector product $\mathcal{T}\mathbf{v}$ only requires a linear complexity of $\mathcal{O}(mn)$. As the action of \mathcal{Y} is a simple reordering of entries which poses virtually no work, computing $\mathcal{Y}\mathcal{T}\mathbf{v}$ needs the same complexity. Alternatively, due to the fact that \mathcal{T} itself is a block Toeplitz matrix, it is well-known that $\mathcal{Y}\mathcal{T}\mathbf{v}$ can be computed in $\mathcal{O}(mn \log n)$ operations using fast Fourier transforms, and the required storage is of $\mathcal{O}(mn)$.

We provide here one way to efficiently invert \mathcal{P} . From (1.6), we can rewrite \mathcal{P} as follows:

$$\begin{aligned}\mathcal{P} &= P_n \otimes (L_m) + 2I_n \otimes I_m \\ &= (\mathbb{S}_n \otimes I_m)(\Lambda_n \otimes L_m + 2I_n \otimes I_m)(\mathbb{S}_n \otimes I_m),\end{aligned}$$

where

$$P_n = \begin{bmatrix} 0 & -1 & & & \\ -1 & 0 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

is a tridiagonal matrix which has the eigendecomposition $P_n = \mathbb{S}_n \Lambda_n \mathbb{S}_n$ with $(\Lambda_n)_{j,j} = -2 \cos\left(\frac{j\pi}{n+1}\right)$ and $\mathbb{S}_n = \sqrt{\frac{2}{n+1}} \left[\sin\left(\frac{ij\pi}{n+1}\right) \right]_{i,j=1}^n \in \mathbb{R}^{n \times n}$ being the discrete sine matrix.

In each iteration of GMRES, it is required to compute $\mathcal{P}^{-1}\mathbf{y}$ for a given vector \mathbf{y} . The product can be implemented via the following three steps:

- Step 1: Compute $\tilde{\mathbf{y}} = (\mathbb{S}_n \otimes I_m)\mathbf{y}$;
- Step 2: Solve $\tilde{\mathbf{z}}^{(k)} = (\lambda_n^{(k)} L_m + 2I_m)^{-1} \tilde{\mathbf{y}}^{(k)}$ for $\tilde{\mathbf{z}}, k = 1, 2, \dots, n$,
where $\Lambda_n = \text{diag}(\lambda_n^{(1)}, \lambda_n^{(2)}, \dots, \lambda_n^{(n)})$ and $\tilde{\mathbf{y}} = [\tilde{\mathbf{y}}^{(1)}; \tilde{\mathbf{y}}^{(2)}; \dots; \tilde{\mathbf{y}}^{(n)}]$;
- Step 3: Compute $\mathbf{z} = (\mathbb{S}_n \otimes I_m)\tilde{\mathbf{z}}$, where $\tilde{\mathbf{z}} = [\tilde{\mathbf{z}}^{(1)}; \tilde{\mathbf{z}}^{(2)}; \dots; \tilde{\mathbf{z}}^{(n)}]$.

Both Steps 1 and 3 can be computed efficiently via DSTs in $\mathcal{O}(mn \log n)$ operations. As for Step 2, the shifted Laplacian systems can be efficiently solved for example using the multigrid methods [9, 29]. For instance, when a usual finite difference scheme with uniform grids is used, the matrix L_m is diagonalizable by a discrete sine matrix. Each system can be solved efficiently by using DSTs in $\mathcal{O}(m \log m)$ operations, which results in a total of $\mathcal{O}(nm \log m)$ operations and $\mathcal{O}(nm)$ storage.

3.1.2. Eigenvalue analysis. For $\mathcal{P}^{-1}\mathcal{Y}\mathcal{T}$, we provide the following theorem:

THEOREM 3.1. *Let $\mathcal{Y}\mathcal{T}, \mathcal{P} \in \mathbb{R}^{mn \times mn}$ be defined by (1.5) and (1.6), respectively. Then, the minimal polynomial of the preconditioned matrix $\mathcal{P}^{-1}\mathcal{Y}\mathcal{T}$ is of at most degree $m + 2$.*

Proof. We first give a decomposition of $\mathcal{M} = \mathcal{P}^{-1}\mathcal{Y}\mathcal{T}$. After direct computations, we have

$$\mathcal{M} = \begin{bmatrix} & & -I_m & O_m \\ & \ddots & \ddots & \\ -I_m & \ddots & & \\ O_m & & & \end{bmatrix}$$

$$\begin{aligned}
 & + \underbrace{\begin{bmatrix} [\mathcal{P}^{-1}]_{1,1} & [\mathcal{P}^{-1}]_{1,2}^T & \cdots & [\mathcal{P}^{-1}]_{1,n}^T \\ [\mathcal{P}^{-1}]_{1,2} & [\mathcal{P}^{-1}]_{2,2} & \cdots & [\mathcal{P}^{-1}]_{2,n}^T \\ \vdots & \vdots & \ddots & \vdots \\ [\mathcal{P}^{-1}]_{1,n} & [\mathcal{P}^{-1}]_{2,n} & \cdots & [\mathcal{P}^{-1}]_{n,n} \end{bmatrix}}_{=\mathcal{P}^{-1}} \begin{bmatrix} O_m & \cdots & O_m & L_m \\ \vdots & & & O_m \\ \vdots & & & \vdots \\ O_m & \cdots & \cdots & O_m \end{bmatrix} \\
 & = \begin{bmatrix} & & -I_m & [\mathcal{P}^{-1}]_{1,1}L_m \\ & \ddots & & \vdots \\ -I_m & & & [\mathcal{P}^{-1}]_{1,n-1}L_m \\ & & & [\mathcal{P}^{-1}]_{1,n}L_m \end{bmatrix} \\
 & = \begin{bmatrix} B & G \\ & [\mathcal{P}^{-1}]_{1,n}L_m \end{bmatrix},
 \end{aligned}$$

where

$$\begin{aligned}
 B & = -Y_{n-1} \otimes I_m \in \mathbb{R}^{m(n-1) \times m(n-1)} \quad \text{and} \\
 G & = [[\mathcal{P}^{-1}]_{1,1}L_m, \dots, [\mathcal{P}^{-1}]_{1,n-1}L_m]^T \in \mathbb{R}^{m(n-1) \times m},
 \end{aligned}$$

with Y_{n-1} being the anti-identity of size $n-1$. Given the block upper-triangular structure of the preconditioned matrix \mathcal{M} , its eigenvalues consist of those of B and of those of $[\mathcal{P}^{-1}]_{1,n}L_m$. In addition, the eigenvalues of $B = -Y_{n-1} \otimes I_m$ are given by $\lambda = 1$ with multiplicity of $m \lfloor \frac{n-1}{2} \rfloor$ and $\lambda = -1$ with multiplicity of $m \lceil \frac{n-1}{2} \rceil$, owing to the tensor structure of B and to the explicit computation of the eigenvalues of Y_{n-1} .

As an intermediate step, $\lambda = 1$ is an eigenvalue of \mathcal{M} with algebraic multiplicity at least $m \lfloor \frac{n-1}{2} \rfloor$, and $\lambda = -1$ is an eigenvalue of \mathcal{M} with algebraic multiplicity at least $m \lceil \frac{n-1}{2} \rceil$. We remark that ‘‘at least’’ refers to the fact that the remaining m eigenvalues of \mathcal{M} , that is, those of $[\mathcal{P}^{-1}]_{1,n}L_m$, may contain again either 1 or -1 .

However, the study of the degree of the minimal polynomial requires the analysis of the geometric multiplicity of the resulting eigenvalues. We consider this matter in detail. Indeed, the matrix B is symmetric and hence diagonalizable so that it has the eigenvalue $\lambda = 1$ with geometric multiplicity $m \lfloor \frac{n-1}{2} \rfloor$ and the eigenvalue $\lambda = -1$ with geometric multiplicity of $m \lceil \frac{n-1}{2} \rceil$. Now, observe that

$$\mathcal{M} \underbrace{\begin{bmatrix} \mathbf{v} \\ 0 \end{bmatrix}}_{=: \mathbf{w}} = \underbrace{\begin{bmatrix} B\mathbf{v} \\ 0 \end{bmatrix}}_{=: \mathbf{z}},$$

where $\mathbf{w}, \mathbf{z} \in \mathbb{R}^{mn}$ and $\mathbf{v} \in \mathbb{R}^{m(n-1)}$. Clearly, if \mathbf{v} is an eigenvector of B , i.e., $B\mathbf{v} = \lambda\mathbf{v}$ with $\lambda = \pm 1$, we have $\mathbf{z} = \lambda\mathbf{w}$ and $\mathcal{M}\mathbf{w} = \lambda\mathbf{w}$. Thus, we have shown that $\lambda = 1$ is an eigenvalue of \mathcal{M} with geometric multiplicity at least $m \lfloor \frac{n-1}{2} \rfloor$ and $\lambda = -1$ is an eigenvalue of \mathcal{M} with geometric multiplicity at least $m \lceil \frac{n-1}{2} \rceil$. Again the wording ‘‘at least’’ refers to the fact that the remaining m eigenvalues of \mathcal{M} may contain again either 1 or -1 .

Let $Q \in \mathbb{R}^{m(n-1) \times m(n-1)}$ be an orthogonal matrix such that $B = QDQ^T$, where

$$D = \begin{bmatrix} I_q & \\ & -I_k \end{bmatrix} \in \mathbb{R}^{m(n-1) \times m(n-1)},$$

with $q = m \lfloor \frac{n-1}{2} \rfloor$ and $k = m \lceil \frac{n-1}{2} \rceil$. By an orthogonal transformation, we know that \mathcal{M} is similar to the matrix $\mathcal{M}^{(1)}$ which is defined as

(3.1)

$$\mathcal{M}^{(1)} = \begin{bmatrix} Q^T & \\ & I_m \end{bmatrix} \mathcal{M} \begin{bmatrix} Q & \\ & I_m \end{bmatrix} = \begin{bmatrix} D & Q^T G \\ & [\mathcal{P}^{-1}]_{1,n} L_m \end{bmatrix} = \begin{bmatrix} I_q & X_m \\ & \vdots \\ -I_k & [\mathcal{P}^{-1}]_{1,n} L_m \end{bmatrix},$$

where X_m is some $m \times m$ matrix. Hence, both $\mathcal{M}^{(1)}$ and \mathcal{M} share the same eigenvalues with the same geometric multiplicity, with the remaining m eigenvalues of \mathcal{M} determined by the block $[\mathcal{P}^{-1}]_{1,n} L_m$.

As a consequence, the minimal polynomial $q_{\mathcal{M}}(t)$ of \mathcal{M} consists of at least two distinct linear factors, namely $q_{\mathcal{M}}(t) = (t-1)(t+1)p(t)$, where $p(t)$ has degree at most equal to m . In other words, $q_{\mathcal{M}}(t)$ has at most degree $m+2$, and this concludes the proof. \square

Theorem 3.1 immediately leads to the following corollary:

COROLLARY 3.2. *The GMRES method for $\mathcal{Y}\mathcal{T}$ with \mathcal{P} as the preconditioner converges in at most $m+2$ iterations in exact arithmetic.*

REMARK 3.3. In general, as mentioned in [41, Chapter 6], the convergence study of preconditioning strategies for nonsymmetric problems is heuristic since descriptive convergence bounds for GMRES or any of the other applicable nonsymmetric Krylov subspace iterative methods do not presently exist. The result with \mathcal{P} in Corollary 3.2 is a non-trivial example whose convergence property associated with GMRES is precisely known.

Before considering the convergence analysis in a MINRES-setting in Section 3.2, we give additional results for the eigenvector-eigenvalue structure of the preconditioned matrix considered in Theorem 3.1 in the subsequent Section 3.1.3.

3.1.3. Eigenstructure. More can be inferred from our proof of Theorem 3.1 on the diagonalizability of \mathcal{M} , which depends on the block $[\mathcal{P}^{-1}]_{1,n} L_m =: Z_m$. Notice that Z_m can be shown to be symmetric according to the discussion in [23, Section 2].

There are the following two scenarios:

- (a) if Z_m does not have ± 1 as its eigenvalues, then \mathcal{M} is diagonalizable;
- (b) if Z_m has ± 1 as its eigenvalues, then \mathcal{M} may be non-diagonalizable in general.

We first consider the case (a): let α be an eigenvalue of Z_m with algebraic multiplicity m_α . Since Z_m is symmetric, we have $\text{rank}(Z_m - \alpha I_m) = m - m_\alpha$. Therefore, from (3.1), we have

$$\text{rank}(\mathcal{M}^{(1)} - \alpha I_{mn}) = k + q + \text{rank}(Z_m - \alpha I_m) = \underbrace{k + q + m}_{=nm} - m_\alpha,$$

since $\alpha \neq \pm 1$. Hence, the geometric multiplicity $\dim(\mathcal{M}^{(1)}) - \text{rank}(\mathcal{M}^{(1)} - \alpha I_{mn})$ is actually the algebraic multiplicity. So $\mathcal{M}^{(1)}$ (and hence \mathcal{M}) is diagonalizable. The proof for this case is finished.

We now show the case (b): let $\alpha = 1$ be an eigenvalue of Z_m with algebraic multiplicity m_α with respect to Z_m and algebraic multiplicity $q + m_\alpha$ with respect to $\mathcal{M}^{(1)}$. Now, if $\tilde{G} = Q^T G$ in (3.1) has full rank and there is a square minor of rank $= \dim(Z_m)$ with rows all having index less or equal to q , then

$$\text{rank}(\mathcal{M}^{(1)} - I_{mn}) = k + \dim(Z_m).$$

Whence the geometric multiplicity with respect to $\mathcal{M}^{(1)}$ is

$\dim(\mathcal{M}^{(1)}) - \text{rank}(\mathcal{M}^{(1)} - I_{mn}) = k + q + \dim(Z_m) - (k + \dim(Z_m)) = q < q + m_\alpha$,

and consequently $\mathcal{M}^{(1)}$ (and hence \mathcal{M}) is not diagonalizable.

More generally, if there is a square minor $G = Q^T G$ of rank $= v < \dim(Z_m)$ with $m_\alpha > \dim(Z_m) - v$ and having rows with index or equal to q , then

$$\text{rank}(\mathcal{M}^{(1)} - I_{mn}) = k + v.$$

Under the previous assumptions the geometric multiplicity with respect to $\mathcal{M}^{(1)}$ is

$$\begin{aligned} \dim(\mathcal{M}^{(1)}) - \text{rank}(\mathcal{M}^{(1)} - I_{mn}) &= k + q + \dim(Z_m) - (k + v) \\ &= q + \dim(Z_m) - v < q + m_\alpha. \end{aligned}$$

Thus, $\mathcal{M}^{(1)}$ (and hence \mathcal{M}) is not diagonalizable. Finally, the same can be shown when $\alpha = -1$ using similar arguments, and the analysis of this case is also concluded.

However, we have $\mathcal{M} = \mathcal{P}^{-1} \mathcal{Y} \mathcal{T}$ with \mathcal{P} symmetric positive definite and $\mathcal{Y} \mathcal{T}$ symmetric and indefinite, so that \mathcal{M} is similar to $\mathcal{P}^{-1/2} \mathcal{Y} \mathcal{T} \mathcal{P}^{-1/2}$ which is symmetric indefinite with the same inertia as $\mathcal{Y} \mathcal{T}$. Hence by the spectral theorem, $\mathcal{P}^{-1/2} \mathcal{Y} \mathcal{T} \mathcal{P}^{-1/2}$ is diagonalizable, and consequently \mathcal{M} is diagonalizable too, showing that the two strict inequalities in the analysis of the case (b) never hold.

3.2. MINRES for $|\mathcal{P}|^{-1} \mathcal{Y} \mathcal{T}$. In this section, we provide a MINRES-approach for $\mathcal{Y} \mathcal{T} \mathbf{u} = \mathcal{Y} \mathbf{f}$. Although \mathcal{P} is symmetric, it is indefinite in general. Therefore, we propose here the use of the absolute-value matrix $|\mathcal{P}| = \sqrt{\mathcal{P}^2}$ as a SPD preconditioner to be used with MINRES.

The matrix-vector product $|\mathcal{P}|^{-1} \mathbf{u}$ for any vector \mathbf{u} can be computed efficiently by simply replacing the eigenvalues in Step 2 of the three-step procedures given in Section 3.1.1 by their absolute value. Before giving the main result for $|\mathcal{P}|$, we first provide the following theorem on matrix functions which will be useful for showing our preconditioning theory.

Let F be analytic on a simply-connected open region of the complex plane containing the interval $[-1, 1]$. Then, there exist an ellipse with foci in -1 and 1 such that F is analytic in their interiors. Let $r_\alpha > 1$ and $r_\beta > 0$ be the half axes of such an ellipse, denoted by $\mathbb{E}_\mathcal{X}$, with $\sqrt{r_\alpha^2 - r_\beta^2} = 1$. The ellipse $\mathbb{E}_\mathcal{X}$ is completely specified by the number $\mathcal{X} := r_\alpha + r_\beta$.

THEOREM 3.4 (Bernstein's theorem [1, Theorem 2.1]). *Let the function F be analytic in the interior of the ellipse $\mathbb{E}_\mathcal{X}$ with $\mathcal{X} > 1$ and continuous on $\mathbb{E}_\mathcal{X}$. In addition, suppose $F(x)$ is real for real x . Then, the best approximation error*

$$E_k(F) := \inf\{\|E - p\|_\infty : \deg(p) \leq k\} \leq \frac{2M(\mathcal{X})}{\mathcal{X}^k(\mathcal{X} - 1)},$$

where $\deg(p)$ denotes the degree of the polynomial $p(x)$ and

$$\|F - p\|_\infty = \max_{-1 \leq x \leq 1} |F(x) - p(x)|, \quad M(\mathcal{X}) = \max_{x \in \mathbb{E}_\mathcal{X}} \{|F(x)|\}.$$

Let A_n be an $n \times n$ symmetric matrix, and let $[\lambda_{\min}, \lambda_{\max}]$ be the smallest interval containing $\sigma(A_n)$. If we introduce the linear affine function

$$\psi(\lambda) = \frac{2\lambda - (\lambda_{\min} + \lambda_{\max})}{\lambda_{\max} - \lambda_{\min}},$$

then $\psi([\lambda_{\min}, \lambda_{\max}]) = [-1, 1]$, and hence the spectrum of the symmetric matrix

$$\psi(A_n) = \frac{2}{\lambda_{\max} - \lambda_{\min}} A_n - \frac{\lambda_{\min} + \lambda_{\max}}{\lambda_{\max} - \lambda_{\min}} I_n$$

is contained in $[-1, 1]$. Provided that a function f analytic on a simply connected region containing $[\lambda_{\min}, \lambda_{\max}]$ and that $f(\lambda)$ is real when λ is real, the function $F = f \circ \psi^{-1}$ satisfies the assumptions of Bernstein's theorem.

In the special case when A_n is SPD and $f(x) = x^{-1/2}$ that we consider in this work, we apply Bernstein's theorem to the function

$$(3.2) \quad F(x) = \frac{1}{\sqrt{\frac{(b-a)}{2}x + \frac{a+b}{2}}},$$

where $a = \lambda_{\min}(A_n)$, $b = \lambda_{\max}(A_n)$, and $1 < \mathcal{X} < \frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}$ with the spectral condition number of A_n being $\kappa = b/a$.

The following proposition indicates that an ideal preconditioner for $\mathcal{Y}\mathcal{T}$ is the SPD preconditioner $\sqrt{\mathcal{T}^T\mathcal{T}}$:

PROPOSITION 3.5. *Let $\mathcal{T} \in \mathbb{R}^{mn \times mn}$ be defined by (1.2). Then $(\sqrt{\mathcal{T}^T\mathcal{T}})^{-1}\mathcal{Y}\mathcal{T}$ has only ± 1 as eigenvalues.*

Proof. Considering the eigendecomposition of the symmetric matrix $\mathcal{Y}\mathcal{T}$, which is $\mathcal{Y}\mathcal{T} = \mathcal{Q}_0^T \Lambda_0 \mathcal{Q}_0$, we have

$$\begin{aligned} (\sqrt{\mathcal{T}^T\mathcal{T}})^{-1}\mathcal{Y}\mathcal{T} &= (\sqrt{(\mathcal{Y}\mathcal{T})^T\mathcal{Y}\mathcal{T}})^{-1}\mathcal{Y}\mathcal{T} = (\sqrt{(\mathcal{Y}\mathcal{T})^2})^{-1}\mathcal{Y}\mathcal{T} \\ &= \mathcal{Q}_0^T (\Lambda_0^2)^{-1/2} \mathcal{Q}_0 \mathcal{Q}_0^T \Lambda_0 \mathcal{Q}_0 = \mathcal{Q}_0^T \tilde{\Lambda}_0 \mathcal{Q}_0, \end{aligned}$$

where $\tilde{\Lambda}_0$ is a diagonal matrix whose entries are either 1 or -1 . Thus, $(\sqrt{\mathcal{T}^T\mathcal{T}})^{-1}\mathcal{Y}\mathcal{T}$ is both symmetric and orthogonal, and hence it has only ± 1 as eigenvalues. \square

In other words, Proposition 3.5 governs the design of an effective preconditioner for $\mathcal{Y}\mathcal{T}$. In what follows, we will show that $|P|$ approximates the ideal preconditioner $\sqrt{\mathcal{T}^T\mathcal{T}}$ in the sense that their difference can be decomposed into a sum of a low-rank matrix and a small norm matrix.

Now, we turn our focus on the following lemma and proposition, which will be used to show our main result.

LEMMA 3.6. *Let $\mathcal{T}, \mathcal{P} \in \mathbb{R}^{mn \times mn}$ be defined in (1.2) and (1.6), respectively. Then,*

$$\text{rank} \left((\mathcal{T}^T\mathcal{T})^K - (\mathcal{P}^2)^K \right) \leq 4Km,$$

for any positive integer K provided that $n > 4Km$.

Proof. Direct computations give

$$\mathcal{T}^T\mathcal{T} - \mathcal{P}^2 = \begin{bmatrix} L_m^2 & & & \\ & -L_m^2 & 2L_m & \\ & 2L_m & -4I_m & \\ & & & \end{bmatrix}.$$

Exploiting such a simple structure of $\mathcal{T}^T \mathcal{T} - \mathcal{P}^2$, we can further show the following result by using a computational lemma given in [35, Lemma 3.11]

$$(\mathcal{T}^T \mathcal{T})^{n_\alpha} (\mathcal{T}^T \mathcal{T} - \mathcal{P}^2) (\mathcal{P}^2)^{n_\beta} = \begin{bmatrix} * & \cdots & * & & & \\ \vdots & & \vdots & & & \\ * & \cdots & * & & & \\ & & & & & \\ & & & & * & \cdots & * \\ & & & & \vdots & & \vdots \\ & & & & * & \cdots & * \end{bmatrix}$$

for integer values n_α and n_β , where $*$ represents a nonzero entry. Namely, the matrix $(\mathcal{T}^T \mathcal{T})^{n_\alpha} (\mathcal{T}^T \mathcal{T} - \mathcal{P}^2) (\mathcal{P}^2)^{n_\beta}$ is a block matrix with two blocks in its Northwest and Southeast corners, respectively, and each block is of size $(n_\alpha + 1)2m \times (n_\beta + 1)2m$. Thus,

$$(\mathcal{T}^T \mathcal{T})^K - (\mathcal{P}^2)^K = \sum_{i=0}^{K-1} (\mathcal{T}^T \mathcal{T})^{K-i-1} (\mathcal{T}^T \mathcal{T} - \mathcal{P}^2) (\mathcal{P}^2)^i$$

is also a block matrix with four blocks in its four corners, and each of them is of size $2Km \times 2Km$, provided that $n > 4Km$. Hence, we have $\text{rank}((\mathcal{T}^T \mathcal{T})^K - (\mathcal{P}^2)^K) \leq 4Km$. \square

PROPOSITION 3.7. *Let $\mathcal{T}, \mathcal{P} \in \mathbb{R}^{mn \times mn}$ be defined in (1.2) and (1.6), respectively. Then, for any $\epsilon > 0$ there exists an integer K such that for all $n > 4Km$*

$$(\sqrt{\mathcal{T}^T \mathcal{T}})^{-1} - |\mathcal{P}|^{-1} = \mathcal{E}_0 + \mathcal{R}_0,$$

where $\|\mathcal{E}_0\|_2 \leq \epsilon$ and $\text{rank}(\mathcal{R}_0) \leq 4Km$. Furthermore, when $n \leq 4Km$, the statement is true with $\mathcal{E}_0 \equiv 0$.

Proof. Let $f(x) = x^{-1/2}$ and $F(x)$ be defined in (3.2). By Theorem 3.4, there exists a polynomial p_K with degree less than or equal to K such that

$$\begin{aligned} \left\| (\sqrt{\mathcal{T}^T \mathcal{T}})^{-1} - p_K(\mathcal{T}^T \mathcal{T}) \right\|_2 &= \left\| (\mathcal{T}^T \mathcal{T})^{-1/2} - p_K(\mathcal{T}^T \mathcal{T}) \right\|_2 \\ &= \max_{x \in \sigma(\mathcal{T}^T \mathcal{T})} |F(x) - p_K(x)| \\ &\leq \|F - p_K(x)\|_\infty \leq \frac{2M(\mathcal{X}_{\mathcal{T}^T \mathcal{T}})}{\mathcal{X}_{\mathcal{T}^T \mathcal{T}} - 1} \cdot \frac{1}{\mathcal{X}_{\mathcal{T}^T \mathcal{T}}^K}, \end{aligned}$$

and

$$\begin{aligned} \left\| |\mathcal{P}|^{-1} - p_K(\mathcal{P}^2) \right\|_2 &= \left\| (\mathcal{P}^2)^{-1/2} - p_K(\mathcal{P}^2) \right\|_2 = \max_{x \in \sigma(\mathcal{P}^2)} |F(x) - p_K(x)| \\ &\leq \|F - p_K(x)\|_\infty \leq \frac{2M(\mathcal{X}_{\mathcal{P}^2})}{\mathcal{X}_{\mathcal{P}^2} - 1} \cdot \frac{1}{\mathcal{X}_{\mathcal{P}^2}^K}, \end{aligned}$$

where

$$1 < \mathcal{X}_{\mathcal{T}^T \mathcal{T}} < \frac{\sqrt{\kappa_{\mathcal{T}^T \mathcal{T}} + 1}}{\sqrt{\kappa_{\mathcal{T}^T \mathcal{T}} - 1}}, \quad 1 < \mathcal{X}_{\mathcal{P}^2} < \frac{\sqrt{\kappa_{\mathcal{P}^2} + 1}}{\sqrt{\kappa_{\mathcal{P}^2} - 1}},$$

and $\kappa_{\mathcal{T}^T \mathcal{T}}$ and $\kappa_{\mathcal{P}^2}$ are the condition numbers of $\mathcal{T}^T \mathcal{T}$ and \mathcal{P}^2 , respectively. Thus, for any $\epsilon > 0$ there exists an integer K such that

$$\left\| (\sqrt{\mathcal{T}^T \mathcal{T}})^{-1} - p_K(\mathcal{T}^T \mathcal{T}) \right\|_2 \leq \epsilon \quad \text{and} \quad \left\| |\mathcal{P}|^{-1} - p_K(\mathcal{P}^2) \right\|_2 \leq \epsilon.$$

Also, we have

$$p_K(\mathcal{T}^T \mathcal{T}) - p_K(\mathcal{P}^2) = \underbrace{\sum_{i=0}^K a_i ((\mathcal{T}^T \mathcal{T})^i - (\mathcal{P}^2)^i)}_{=: \mathcal{R}_0}.$$

By Lemma 3.6, we know that \mathcal{R}_0 has the same sparsity structure as that of $(\mathcal{T}^T \mathcal{T})^K - (\mathcal{P}^2)^K$. Consequently, we deduce that $\text{rank}(\mathcal{R}_0) \leq 4Km$.

We then obtain

$$\begin{aligned} & (\sqrt{\mathcal{T}^T \mathcal{T}})^{-1} - |\mathcal{P}|^{-1} \\ &= \underbrace{(\sqrt{\mathcal{T}^T \mathcal{T}})^{-1} - p_K(\mathcal{T}^T \mathcal{T}) + p_K(\mathcal{P}^2)}_{=: \mathcal{E}_0} - |\mathcal{P}|^{-1} + \underbrace{p_K(\mathcal{T}^T \mathcal{T}) - p_K(\mathcal{P}^2)}_{=: \mathcal{R}_0}, \end{aligned}$$

where $\|\mathcal{E}_0\|_2 \leq 2\epsilon$ and $\text{rank}(\mathcal{R}_0) \leq 4Km$. Therefore, the proof is concluded. \square

At last, we are ready to provide the following theorem which guarantees the effectiveness of $|\mathcal{P}|$.

THEOREM 3.8. *Let $\mathcal{Y}\mathcal{T}, \mathcal{P} \in \mathbb{R}^{mn \times mn}$ be defined by (1.5) and (1.6), respectively. Then, for any $\epsilon > 0$ there exists an integer K such that for all $n > 4Km$*

$$|\mathcal{P}|^{-1} \mathcal{Y}\mathcal{T} = \mathcal{Q} + \mathcal{E} + \mathcal{R},$$

where \mathcal{Q} is both symmetric and orthogonal, $\|\mathcal{E}\|_2 \leq \epsilon$, and $\text{rank}(\mathcal{R}) \leq 4Km$.

Proof. By Proposition 3.7, we have

$$\begin{aligned} |\mathcal{P}|^{-1} \mathcal{Y}\mathcal{T} &= \left((\sqrt{\mathcal{T}^T \mathcal{T}})^{-1} - \mathcal{E}_0 - \mathcal{R}_0 \right) \mathcal{Y}\mathcal{T} \\ &= \underbrace{(\sqrt{\mathcal{T}^T \mathcal{T}})^{-1} \mathcal{Y}\mathcal{T}}_{=: \mathcal{Q}} + \underbrace{(-\mathcal{E}_0 \mathcal{Y}\mathcal{T})}_{=: \mathcal{E}} + \underbrace{(-\mathcal{R}_0 \mathcal{Y}\mathcal{T})}_{=: \mathcal{R}}, \end{aligned}$$

where \mathcal{Q} is both symmetric and orthogonal by Proposition 3.5,

$$\text{rank}(\mathcal{R}) \leq \text{rank}(\mathcal{R}_0 \mathcal{Y}\mathcal{T}) \leq 4Km \quad \text{and} \quad \|\mathcal{E}\|_2 = \|-\mathcal{E}_0 \mathcal{Y}\mathcal{T}\|_2 \leq \|\mathcal{Y}\mathcal{T}\|_2 \epsilon.$$

Lastly, we have that

$$\|\mathcal{Y}\mathcal{T}\|_2 = \|\mathcal{T}\|_2 \leq \|g\|_\infty$$

by using the general inequality in [37, Corollary 4.2] where the Schatten norm with $p = \infty$ is equivalent to the spectral norm $\|\cdot\|_2$. Hence, $\|\mathcal{Y}\mathcal{T}\|_2$ is uniformly bounded with respect to n , and the proof is concluded. \square

As a consequence of Theorem 3.8 and [3, Corollary 3], we know that for large enough n , the preconditioned matrix sequence $\{|\mathcal{P}|^{-1} \mathcal{Y}\mathcal{T}\}_n$ has clustered eigenvalues around ± 1 , with a number of outliers independent of n . Hence, the convergence is independent of the time step, and we can expect that MINRES for $\mathcal{Y}\mathcal{T}$ will converge rapidly in exact arithmetic with $|\mathcal{P}|$ as the preconditioner.

3.3. CGNE for $\mathcal{P}^{-1} \mathcal{T}$. For a complete analysis, we consider in this section a CGNE method, even though we do not mainly use normal equations in the numerical examples provided in Section 4.

It is worth noticing that

$$(\mathcal{P}^{-1}\mathcal{Y}\mathcal{T})^T\mathcal{P}^{-1}\mathcal{Y}\mathcal{T} = \mathcal{T}^T\mathcal{Y}\mathcal{P}^{-2}\mathcal{Y}\mathcal{T} = \mathcal{T}^T\mathcal{P}^{-2}\mathcal{T} = (\mathcal{P}^{-1}\mathcal{T})^T\mathcal{P}^{-1}\mathcal{T}.$$

Hence, the above considered normal equation systems are equivalent, and it suffices to only consider the eigenvalues of $(\mathcal{P}^{-1}\mathcal{T})^T\mathcal{P}^{-1}\mathcal{T}$.

The following result guarantees a convergence rate which is independent of the time-steps.

THEOREM 3.9. *Let $\mathcal{T}, \mathcal{P} \in \mathbb{R}^{mn \times mn}$ be defined by (1.2) and (1.6), respectively. Then, the minimal polynomial of the preconditioned matrix $(\mathcal{P}^{-1}\mathcal{T})^T\mathcal{P}^{-1}\mathcal{T}$ is of at most degree $2m + 1$.*

Proof.

We first let $\mathcal{M} = (\mathcal{P}^{-1}\mathcal{T})^T\mathcal{P}^{-1}\mathcal{T}$ and compute

$$\mathcal{M} = \begin{bmatrix} I_m & & & -[\mathcal{P}^{-1}]_{1,n-1}L_m \\ & \ddots & & \vdots \\ & & I_m & [-\mathcal{P}^{-1}]_{1,1}L_m \\ -([\mathcal{P}^{-1}]_{1,n-1}L_m)^T & \cdots & -([\mathcal{P}^{-1}]_{1,1}L_m)^T & ([\mathcal{P}^{-1}]_{1,n}L_m)^T([\mathcal{P}^{-1}]_{1,n}L_m) \end{bmatrix}.$$

Similar to proving Theorem 3.9, it can be shown that $\lambda = 1$ has geometric multiplicity of $m(n - 2)$. Thus, the minimal polynomial $q_{\mathcal{M}}(t)$ of \mathcal{M} consists of one distinct linear factor of order 1, namely $q_{\mathcal{M}}(t) = (t - 1)p(t)$, where $p(t)$ has degree of at most $2m$. In other words, $q_{\mathcal{M}}(t)$ has at most degree $2m + 1$. \square

COROLLARY 3.10. *The CGNE method for \mathcal{T} with \mathcal{P} as the preconditioner converges in at most $2m + 1$ iterations in exact arithmetic.*

4. Numerical examples. In this section, we demonstrate the effectiveness of our proposed solvers. All numerical experiments are carried out using MATLAB on a HP EliteDesk 800 G5 Small Form Factor PC with Intel Core i7-9700 CPU @ 3.00GHz with 16GB RAM. The CPU time in seconds is measured using the built-in function `tic/toc`. All Krylov subspace solvers are implemented using the build-in functions in MATLAB. Furthermore, we choose a zero initial guess and a stopping tolerance of 10^{-6} based on the reduction in the relative residual norms unless indicated otherwise.

Throughout all examples, we consider finite difference methods with uniform spatial grids, which results in K_m being diagonalized by the discrete sine transform. Notice that in the tables below ‘‘DoF’’ denotes the degree of freedom, and the existing block circulant preconditioner \mathcal{S} (or its SPD counterpart $|\mathcal{S}| := \sqrt{\mathcal{S}^T\mathcal{S}}$) denoted by (1.4) is compared. Also, left-preconditioned GMRES is used for all tests. When \mathcal{S} is used with GMRES, we indicate that the concerned preconditioned matrix is $\mathcal{S}^{-1}\mathcal{T}$, i.e., \mathcal{Y} is not applied to the original matrix \mathcal{T} . That is in contrast with the use of \mathcal{P} ; the preconditioned matrix is $\mathcal{P}^{-1}\mathcal{Y}\mathcal{T}$ when \mathcal{P} is used with GMRES. As for the absolute-value preconditioner, the concerned matrix is $|\mathcal{P}|^{-1}\mathcal{Y}\mathcal{T}$ (or $|\mathcal{S}|^{-1}\mathcal{Y}\mathcal{T}$) when $|\mathcal{P}|$ (or $|\mathcal{S}|$) is used with MINRES. The discretization scheme provided in (1.2) and the simple central difference scheme in (1.3) are denoted by Schemes I and II, respectively.

EXAMPLE 4.1. The first example is [6, Section 3.1.2]. We illustrate our proposed solvers by considering (1.1) in the unilevel case, which corresponds to the following linear ordinary differential equation:

$$\begin{cases} u_{tt} = au + f, & t \in (0, T], \\ u(0) = \psi_0, \\ u_t(0) = \psi_1. \end{cases}$$

where f , ψ_0 , and ψ_1 are given functions. Since the block size is $m = 1$ for this example, the matrix in (1.5) reduces to the symmetric Toeplitz matrix

$$\mathcal{Y}\mathcal{T} = \begin{bmatrix} & & & 1 - a\frac{\tau^2}{2} & -2 & 1 - a\frac{\tau^2}{2} \\ & & \ddots & \ddots & \ddots & \\ & 1 - a\frac{\tau^2}{2} & & -2 & 1 - a\frac{\tau^2}{2} & \\ & -2 & & 1 - a\frac{\tau^2}{2} & & \\ 1 - a\frac{\tau^2}{2} & & & & & \end{bmatrix},$$

and our proposed preconditioner is

$$\mathcal{P} = \begin{bmatrix} 2 & -(1 - a\frac{\tau^2}{2}) & & & & \\ -(1 - a\frac{\tau^2}{2}) & 2 & -(1 - a\frac{\tau^2}{2}) & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & -(1 - a\frac{\tau^2}{2}) & 2 & \\ & & & & 2 & \end{bmatrix}.$$

Table 4.1 displays the iteration numbers for solving $\mathcal{Y}\mathcal{T}$ with $T = 10^3$, $a = -1$, $f = 0$, $\psi_0 = 1$, and $\psi_1 = -1$. When GMRES is applied, we observe that our proposed preconditioner \mathcal{P} is comparable with the existing preconditioner \mathcal{S} . When CGNE is used, our main preconditioner \mathcal{P} requires fewer iterations compared with \mathcal{S} . Also, the iteration numbers agree excellently with our theorems.

Notice that the CPU times are not reported for this example since they are too small to display any significant difference for comparison.

TABLE 4.1
Iteration numbers with GMRES and CGNE for Example 4.1.

n	GMRES		CGNE	
	\mathcal{S}	\mathcal{P}	\mathcal{S}	\mathcal{P}
2^{12}	3	3	6	3
2^{13}	3	3	6	3
2^{14}	3	3	6	3
2^{15}	3	3	6	3

EXAMPLE 4.2. The second example is tested in [18]. This is a one-dimensional problem with $\Omega = (0, 1)$ and $\psi_0(x) = \chi_{[3/8, 5/8]}(x) \cos^2(4\pi(x - 1/2))$, $\psi_1(x) = 0$, and $f = 0$, where χ_X is the indicator function on the set X .

One can observe from Table 4.2 the negative result that the GMRES iteration numbers grow with n and m when Scheme I is used. We however emphasize that our new preconditioner \mathcal{P} does improve convergence, outmatching \mathcal{S} by consistently requiring fewer iterations and shorter CPU times. We report similar convergence results in Table 4.3 when Scheme II is used, except when $n = m$. In those cases, we observe mesh-independent convergence with \mathcal{P} , while \mathcal{S} fails to be working due to being nearly singular.

TABLE 4.2
Convergence results with GMRES for Example 4.2 (Scheme I).

n	GMRES		\mathcal{S}		\mathcal{P}	
	$m + 1$	DoF	Iter	CPU	Iter	CPU
2^8	2^5	7936	31	$5.2 \cdot 10^{-2}$	9	$2.4 \cdot 10^{-2}$
	2^6	16128	47	$1.6 \cdot 10^{-1}$	22	$7.5 \cdot 10^{-2}$
	2^7	32512	68	$7.3 \cdot 10^{-1}$	37	$2.1 \cdot 10^{-1}$
	2^8	65280	94	$2.2 \cdot 10^0$	53	$5.7 \cdot 10^{-1}$
2^9	2^6	32256	50	$2.8 \cdot 10^{-1}$	26	$1.3 \cdot 10^{-1}$
	2^7	65024	67	$1.2 \cdot 10^0$	36	$3.3 \cdot 10^{-1}$
	2^8	130560	95	$4.4 \cdot 10^0$	50	$1.1 \cdot 10^0$
	2^9	261632	128	$1.4 \cdot 10^1$	93	$5.5 \cdot 10^0$
2^{10}	2^7	130048	66	$2.5 \cdot 10^0$	28	$5.2 \cdot 10^{-1}$
	2^8	261120	86	$8.1 \cdot 10^0$	46	$2.0 \cdot 10^0$
	2^9	523264	142	$4.6 \cdot 10^1$	68	$6.8 \cdot 10^0$
	2^{10}	1047552	180	$1.7 \cdot 10^2$	104	$3.3 \cdot 10^1$
2^{11}	2^8	522240	91	$2.2 \cdot 10^1$	37	$3.4 \cdot 10^0$
	2^9	1046528	137	$1.1 \cdot 10^2$	45	$1.2 \cdot 10^1$
	2^{10}	2095104	190	$4.0 \cdot 10^2$	88	$5.8 \cdot 10^1$
	2^{11}	4192256	-	-	166	$4.1 \cdot 10^2$

TABLE 4.3
Convergence results with GMRES for Example 4.2 (Scheme II).

n	GMRES		\mathcal{S}		\mathcal{P}	
	$m + 1$	DoF	Iter	CPU	Iter	CPU
2^8	2^5	7936	33	$5.1 \cdot 10^{-2}$	14	$2.5 \cdot 10^{-2}$
	2^6	16128	43	$1.2 \cdot 10^{-1}$	22	$7.2 \cdot 10^{-2}$
	2^7	32512	64	$6.4 \cdot 10^{-1}$	34	$1.9 \cdot 10^{-1}$
	2^8	65280	-	-	3	$3.9 \cdot 10^{-2}$
2^9	2^6	32256	45	$2.4 \cdot 10^{-1}$	23	$1.0 \cdot 10^{-1}$
	2^7	65024	72	$1.4 \cdot 10^0$	32	$2.7 \cdot 10^{-1}$
	2^8	130560	99	$4.6 \cdot 10^0$	52	$1.2 \cdot 10^0$
	2^9	261632	-	-	3	$1.4 \cdot 10^{-1}$
2^{10}	2^7	130048	61	$2.3 \cdot 10^0$	32	$6.2 \cdot 10^{-1}$
	2^8	261120	96	$9.5 \cdot 10^0$	39	$1.8 \cdot 10^0$
	2^9	523264	134	$4.1 \cdot 10^1$	64	$6.3 \cdot 10^0$
	2^{10}	1047552	-	-	4	$7.9 \cdot 10^{-1}$
2^{11}	2^8	522240	91	$3.0 \cdot 10^1$	29	$3.0 \cdot 10^0$
	2^9	1046528	130	$1.2 \cdot 10^2$	53	$1.6 \cdot 10^1$
	2^{10}	2095104	187	$3.7 \cdot 10^2$	90	$5.8 \cdot 10^1$
	2^{11}	4192256	-	-	4	$3.9 \cdot 10^0$

EXAMPLE 4.3. In the third example, we consider a two-dimensional wave equation defined in $\Omega = (0, 1)^2$ with

$$T = 1, u_0(x, y) = \sin(\pi x) \sin(\pi y), u_1(x, y) = 3 \sin(\pi x) \sin(\pi y)$$

and

$$f(x, y, t) = 6 \sin(\pi x) \sin(\pi y)(t + 1) + 2\pi^2 \sin(\pi x) \sin(\pi y)(t + 1)^3.$$

The exact solution is $u(x, y, t) = \sin(\pi x) \sin(\pi y)(t + 1)^3$.

Table 4.4 displays the convergence results when Scheme I is used. While we observe similar convergence behaviors with \mathcal{S} and \mathcal{P} (or $|\mathcal{S}|$ and $|\mathcal{P}|$), our preconditioner leads to faster convergence in most cases. For this example, we report that the GMRES solver appears to be best compared with MINRES. The latter approach might have room for improvements due to the growing iterations numbers induced. Therefore, we recommend the use of GMRES for this example and stress that advancing the MINRES solver for wave equations is worth further investigation, which is left for future work.

TABLE 4.4
 Convergence results with GMRES and MINRES for Example 4.3 (Scheme I).

GMRES			\mathcal{S}		\mathcal{P}	
n	$m + 1$	DoF	Iter	CPU	Iter	CPU
2^6	2^3	3136	3	$2.2 \cdot 10^{-2}$	3	$1.8 \cdot 10^{-2}$
	2^4	14400	3	$2.9 \cdot 10^{-2}$	3	$2.5 \cdot 10^{-2}$
	2^5	61504	3	$1.0 \cdot 10^{-1}$	3	$5.9 \cdot 10^{-2}$
	2^6	254016	3	$2.2 \cdot 10^{-1}$	4	$2.2 \cdot 10^{-1}$
2^7	2^4	28800	3	$4.7 \cdot 10^{-2}$	3	$4.0 \cdot 10^{-2}$
	2^5	123008	3	$2.0 \cdot 10^{-1}$	3	$9.8 \cdot 10^{-2}$
	2^6	508032	3	$7.0 \cdot 10^{-1}$	4	$4.6 \cdot 10^{-1}$
	2^7	2064512	3	$2.6 \cdot 10^0$	6	$2.3 \cdot 10^0$
2^8	2^5	522240	3	$4.2 \cdot 10^{-1}$	3	$2.4 \cdot 10^{-1}$
	2^6	1016064	3	$1.5 \cdot 10^0$	3	$8.5 \cdot 10^{-1}$
	2^7	4129024	9	$1.3 \cdot 10^1$	6	$5.3 \cdot 10^0$
	2^8	16646400	10	$5.8 \cdot 10^1$	15	$5.5 \cdot 10^1$
MINRES			$ \mathcal{S} $		$ \mathcal{P} $	
n	$m + 1$	DoF	Iter	CPU	Iter	CPU
2^6	2^3	3136	6	$3.3 \cdot 10^{-2}$	6	$3.0 \cdot 10^{-2}$
	2^4	14400	6	$5.9 \cdot 10^{-2}$	5	$3.5 \cdot 10^{-2}$
	2^5	61504	12	$2.1 \cdot 10^{-1}$	6	$5.9 \cdot 10^{-2}$
	2^6	254016	14	$6.8 \cdot 10^{-1}$	14	$6.2 \cdot 10^{-1}$
2^7	2^4	28800	6	$9.4 \cdot 10^{-2}$	5	$5.8 \cdot 10^{-2}$
	2^5	123008	12	$4.0 \cdot 10^{-1}$	6	$1.7 \cdot 10^{-1}$
	2^6	508032	18	$1.8 \cdot 10^0$	10	$9.8 \cdot 10^{-1}$
	2^7	2064512	30	$1.1 \cdot 10^1$	27	$8.2 \cdot 10^0$
2^8	2^5	522240	8	$5.8 \cdot 10^{-1}$	6	$4.0 \cdot 10^{-1}$
	2^6	1016064	18	$3.7 \cdot 10^0$	10	$2.1 \cdot 10^0$
	2^7	4129024	56	$4.2 \cdot 10^1$	24	$1.6 \cdot 10^1$
	2^8	16646400	84	$3.2 \cdot 10^2$	90	$2.7 \cdot 10^2$

5. Conclusions. We have proposed a novel block Toeplitz preconditioner \mathcal{P} (or $|\mathcal{P}|$) for the modified all-at-once system of wave equations $\mathcal{Y}\mathcal{T}\mathbf{u} = \mathcal{Y}\mathbf{f}$, and we have shown its efficiency in the construction and the preconditioning effect both via a detailed theoretical

study and numerical evidence. Our preconditioner is designed according to the previous work for the spectral distribution of symmetrized Toeplitz matrix sequences [11, 31]. As shown in the numerical examples, our preconditioner in most cases can outperform the existing block circulant preconditioner \mathcal{S} in both iteration numbers and CPU times, while having the same parallel-in-time character. Our method provides an alternative preconditioning approach based on DSTs, which advances the symmetrization preconditioning theory pioneered in [33] whose development is still at its infancy.

Concerning a potential future work, we will integrate our proposed methodology with the α -block circulant preconditioning technique [25, 28, 39] in order to further generalize the adopted symmetrization solver with MINRES for wave equations. Another direction for future research is to develop effective preconditioners based on our approach for wave equations with time-varying coefficients and nonlinear wave equations. In both directions, the use of the spectral tools from the GLT technology [17] will be beneficial.

Finally, we remark that the use of the eigenvalue/eigenvector results provided in the theoretical analysis and especially in Section 3.1.3 could be the key for understanding the reason why the number of actual observed iterations is sometimes much lower than that predicted in our derivations in Section 3.

Acknowledgments. The authors would like to thank the anonymous referees for their careful reading and helpful suggestions. The work of Sean Hon was supported in part by the Hong Kong RGC under grant 22300921, a start-up grant from the Croucher Foundation, and a Tier 2 Start-up Grant from Hong Kong Baptist University. The work of Stefano Serra-Capizzano was supported in part by INDAM-GNCS and was funded from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 955701. The JU receives support from the European Union’s Horizon 2020 research and innovation programme and Belgium, France, Germany, Switzerland.

REFERENCES

- [1] M. BENZI AND G. H. GOLUB, *Bounds for the entries of matrix functions with applications to preconditioning*, BIT Numer. Math., 39 (1999), pp. 417–438.
- [2] D. BINI AND F. DI BENEDETTO, *A new preconditioner for the parallel solution of positive definite Toeplitz systems*, in Proc. Second ACM Symp. on Parallel Algorithms and Architectures, ACM, New York, 1990, pp. 220–223.
- [3] J. H. BRANDTS AND R. REIS DA SILVA, *Computable eigenvalue bounds for rank- k perturbations*, Linear Algebra Appl., 432 (2010), pp. 3100–3116.
- [4] R. H. CHAN AND X.-Q. JIN, *An Introduction to Iterative Toeplitz Solvers*, SIAM, Philadelphia, 2007.
- [5] R. H. CHAN AND M. K. NG, *Conjugate gradient methods for Toeplitz systems*, SIAM Rev., 38 (1996), pp. 427–482.
- [6] F. DANIELI AND A. J. WATHEN, *All-at-once solution of linear wave equations*, Numer. Linear Algebra Appl., 28 (2021), Art. e2386, 16 pages.
- [7] F. DI BENEDETTO AND S. SERRA-CAPIZZANO, *A unifying approach to abstract matrix algebra preconditioning*, Numer. Math., 82 (1999), pp. 57–90.
- [8] V. A. DOBREV, TZ. KOLEV, N. A. PETERSSON, AND J. B. SCHRODER, *Two-level convergence theory for multigrid reduction in time (MGRIT)*, SIAM J. Sci. Comput., 39 (2017), pp. S501–S527.
- [9] Y. A. ERLANGGA, C. W. OOSTERLEE, AND C. VUIK, *A novel multigrid based preconditioner for heterogeneous Helmholtz problems*, SIAM J. Sci. Comput., 27 (2006), pp. 1471–1492.
- [10] R. D. FALGOUT, S. FRIEDHOFF, TZ. V. KOLEV, S. P. MACLACHLAN, AND J. B. SCHRODER, *Parallel time integration with multigrid*, SIAM J. Sci. Comput., 36 (2014), pp. C635–C661.
- [11] P. FERRARI, I. FURCI, S. HON, M. A. MURSALEEN, AND S. SERRA-CAPIZZANO, *The eigenvalue distribution of special 2-by-2 block matrix-sequences with applications to the case of symmetrized Toeplitz structures*, SIAM J. Matrix Anal. Appl., 40 (2019), pp. 1066–1086.
- [12] M. J. GANDER, *50 years of time parallel time integration*, in Multiple Shooting and Time Domain Decomposition Methods, T. Carraro, M. Geiger, S. Körkel, and R. Rannacher, eds., Contrib. Math. Comput. Sci., 9, Springer, Cham, 2015, pp. 69–113.

- [13] M. J. GANDER, L. HALPERN, J. RYAN, AND T. T. B. TRAN, *A direct solver for time parallelization*, in Domain Decomposition Methods in Science and Engineering XXII, T. Dickopf, M. J. Gander, L. Halpern, R. Krause, and L. F. Pavarino, eds., Lect. Notes Comput. Sci. Eng., 104, Springer, Cham, 2016, pp. 491–499.
- [14] M. J. GANDER, J. LIU, S. L. WU, X. YUE, AND T. ZHOU, *ParaDiag: parallel-in-time algorithms based on the diagonalization technique*, Preprint on arXiv, 2020. <https://arxiv.org/abs/2005.09158>
- [15] M. J. GANDER AND M. NEUMÜLLER, *Analysis of a new space-time parallel multigrid algorithm for parabolic problems*, SIAM J. Sci. Comput. 38 (2016), pp. A2173–A2208.
- [16] M. J. GANDER AND S. VANDEWALLE, *Analysis of the parareal time-parallel time-integration method*, SIAM J. Sci. Comput., 29 (2007), pp. 556–578.
- [17] C. GARONI AND S. SERRA-CAPIZZANO, *Generalized Locally Toeplitz Sequences: Theory and Applications. Vol 2*, Springer, Cham, 2018.
- [18] A. GODDARD AND A. J. WATHEN, *A note on parallel preconditioning for all-at-once evolutionary PDEs*, Electron. Trans. Numer. Anal., 51 (2019), pp. 135–150.
<http://etna.ricam.oeaw.ac.at/vol.51.2019/pp135-150.dir/pp135-150.pdf>
- [19] S. HON AND S. SERRA-CAPIZZANO, *A sine transform based preconditioned MINRES method for all-at-once systems from evolutionary partial differential equations*, Preprint on arXiv, 2022.
<https://arxiv.org/abs/2201.10062>
- [20] S. HON, S. SERRA-CAPIZZANO, AND A. J. WATHEN, *Band-Toeplitz preconditioners for ill-conditioned Toeplitz systems*, BIT Numer. Math., 62 (2022), pp. 465–491.
- [21] G. HORTON AND S. VANDEWALLE, *A space-time multigrid method for parabolic partial differential equations*, SIAM J. Sci. Comput., 16 (1995), pp. 848–864.
- [22] X.-Q. JIN, *Developments and Applications of Block Toeplitz Iterative Solvers*, Kluwer, Dordrecht; Science Press Beijing, Beijing, 2002.
- [23] J. JAIN, H. LI, S. CAULEY, C.-K. KOH, AND V. BALAKRISHNAN, *Numerically stable algorithms for inversion of block tridiagonal and banded matrices*, ECE Technical Reports, Paper 357, Purdue University, West Lafayette, 2007.
- [24] B. LI, J. LIU, AND M. XIAO, *A fast and stable preconditioned iterative method for optimal control problem of wave equations*, SIAM J. Sci. Comput., 37 (2015), pp. A2508–A2534.
- [25] X.-L. LIN AND M. K. NG, *An all-at-once preconditioner for evolutionary partial differential equations*, SIAM J. Sci. Comput., 43 (2021), pp. A2766–A2784.
- [26] X.-L. LIN, M. K. NG, AND Y. ZHI, *A parallel-in-time two-sided preconditioning for all-at-once system from a non-local evolutionary equation with weakly singular kernel*, J. Comput. Phys., 434 (2021), Art. 110221, 13 pages.
- [27] J. L. LIONS, Y. MADAY, AND G. TURINICI, *“parareal” in time discretization of PDE’s*, C. R. Acad. Sci. Paris Sér. I Math., 332 (2001), pp. 661–668.
- [28] J. LIU AND S.-L. WU, *A fast block α -circulant preconditioner for all-at-once systems from wave equations*, SIAM J. Matrix Anal. Appl., 41 (2020), pp. 1912–1943.
- [29] S. P. MACLACHLAN AND C. W. OOSTERLEE, *Algebraic multigrid solvers for complex-valued matrices*, SIAM J. Sci. Comput., 30 (2008), pp. 1548–1571.
- [30] Y. MADAY AND E. M. RÖNQUIST, *Parallelization in time through tensor-product space–time solvers*, C. R. Math. Acad. Sci. Paris, 346 (2008), pp. 113–118.
- [31] M. MAZZA AND J. PESTANA, *Spectral properties of flipped Toeplitz matrices and related preconditioning*, BIT Numer. Math., 59 (2019), pp. 463–482.
- [32] E. McDONALD, S. HON, J. PESTANA, AND A. J. WATHEN, *Preconditioning for nonsymmetry and time-dependence*, in Domain Decomposition Methods in Science and Engineering XXIII, C.-O. Lee, X.-C. Cai, D. E. Keyes, H. H. Kim, A. Klawonn, E.-J. Park, and O. B. Widlund, eds., Lect. Notes Comput. Sci. Eng., 116, Springer, Cham, 2017, pp. 81–91.
- [33] E. McDONALD, J. PESTANA, AND A. J. WATHEN, *Preconditioning and iterative solution of all-at-once systems for evolutionary partial differential equations*, SIAM J. Sci. Comput., 40 (2018), pp. A1012–A1033.
- [34] M. K. NG, *Iterative Methods for Toeplitz Systems*, Oxford University Press, New York, 2004.
- [35] M. K. NG AND J. PAN, *Approximate inverse circulant-plus-diagonal preconditioners for Toeplitz-plus-diagonal matrices*, SIAM J. Sci. Comput., 32 (2010), pp. 1442–1464.
- [36] S. SERRA, *Asymptotic results on the spectra of block Toeplitz preconditioned matrices*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 31–44.
- [37] S. SERRA AND P. TILLI, *On unitarily invariant norms of matrix-valued linear positive operators*, J. Inequal. Appl., 7 (2002), pp. 309–330.
- [38] S. SERRA-CAPIZZANO, *Toeplitz preconditioners constructed from linear approximation processes*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 446–465.
- [39] Y. SUN, S. L. WU, AND Y. XU, *A parallel-in-time implementation of the Numerov method for wave equations*, J. Sci. Comput., 90 (2022), Art. 20, 31 pages.

- [40] P. TILLI, *A note on the spectral distribution of Toeplitz matrices*, Linear and Multilinear Algebra, 45 (1998), pp. 147–159.
- [41] A. J. WATHEN, *Preconditioning*, Acta Numer., 24 (2015), pp. 329–376.
- [42] S. L. WU AND T. ZHOU, *Parallel implementation for the two-stage SDIRK methods via diagonalization*, J. Comput. Phys., 428 (2021), Art. 110076, 18 pages.