BIT manuscript No.
(will be inserted by the editor)

Band-Toeplitz preconditioners for ill-conditioned Toeplitz systems

Sean Hon · Stefano Serra-Capizzano · Andy Wathen

Received: date / Accepted: date

Abstract Preconditioning for Toeplitz systems has been an active research area over the past few decades. Along this line of research, circulant preconditioners have been recently proposed for the Toeplitz-like system arising from discretizing fractional diffusion equations. A common approach is to combine a circulant preconditioner with the preconditioned conjugate gradient normal residual (PCGRN) method for the coefficient system. In this work, instead of using PCGRN for the normal equation system, we propose a simple yet effective preconditioning approach for solving the original system using the preconditioned minimal residual (PMINRES) method that can achieve convergence guarantees depending only on eigenvalues. Namely, for a large class of ill-conditioned Toeplitz systems, we propose a number of preconditioners that attain the overall optimal $\mathcal{O}(n \log n)$ complexity. We first symmetrize the given Toeplitz system by using a permutation matrix and construct a band-Toeplitz plus circulant preconditioner for the modified system. Then, under certain assumptions, we show that the eigenvalues of the preconditioned system are clustered around ± 1 except a number of outliers and hence superlinear convergence rate of PMINRES can be achieved. Particularly, we indicate that

The work of Hon was supported in part by the Hong Kong RGC under grant 22300921 and a start-up allowance from the Croucher Foundation.

Sean Hon
Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong
E-mail: seanyshon@hkbu.edu.hk

Stefano Serra-Capizzano
Department of Science and High Technology, University of Insubria, 22100, Como, Italy and Department of Information Technology, Division of Scientific Computing, Uppsala University, ITC, SE-751 05, Uppsala, Sweden
E-mail: stefano.serrac@uninsubria.it

Andy Wathen
Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Oxford, OX2 6GG, United Kingdom
E-mail: wathen@maths.ox.ac.uk

our solver can be applied to solve certain fractional diffusion equations. An extension of this work to the block Toeplitz case is also included. Numerical examples are provided to demonstrate the effectiveness of our proposed method.

Keywords Toeplitz/Hankel matrices · band-Toeplitz/circulant preconditioners · fractional diffusion equations · singular value/eigenvalue distribution · Krylov subspace methods · block matrices

Mathematics Subject Classification (2000) 15B05 · 35R11 · 65F08 · 65F10 · 65M22

1 Introduction

Over the past few decades, developing effective numerical methods for Toeplitz systems has been a research focus. One example of applications that has brought much attention in the recent years is solving fractional diffusion equations (FDEs). In particular, various circulant preconditioners have been proposed for the Toeplitz coefficient system arising from discretizing FDEs since [27]. For simplicity, we consider the following one-dimensional initial-boundary value problem and solve it for $u(x,t)$

$$\begin{cases} \frac{\partial u(x,t)}{\partial t} = d_+ \frac{\partial^\alpha u(x,t)}{\partial_+ x^\alpha} + d_- \frac{\partial^\alpha u(x,t)}{\partial_- x^\alpha} + f(x,t), & (x,t) \in (a,b)\times(0,T], \\ u(a,t) = u(b,t) = 0, & t \in (0,T], \\ u(x,0) = u_0(x), & x \in (a,b), \end{cases} \tag{1.1}$$

where $\alpha \in (1,2)$ is the fractional derivative order, $f(x,t)$ is the source term, and $d_\pm$ are the nonnegative diffusion coefficients. The Riemann-Liouville fractional derivatives are defined as

$$\begin{aligned} \frac{\partial^\alpha u(x,t)}{\partial_+ x^\alpha} &= \frac{1}{\Gamma(n-\alpha)} \frac{\partial^n}{\partial x^n} \int_0^x \frac{u(\xi,t)}{(x-\xi)^{\alpha+1-n}}\, d\xi, \\ \frac{\partial^\alpha u(x,t)}{\partial_- x^\alpha} &= \frac{(-1)^n}{\Gamma(n-\alpha)} \frac{\partial^n}{\partial x^n} \int_x^0 \frac{u(\xi,t)}{(\xi-x)^{\alpha+1-n}}\, d\xi, \end{aligned}$$

where Γ is the gamma function and n is the integer such that $n-1 < \alpha \le n$. When α is an integer, the fractional derivatives reduce to the usual derivatives.

We can partition the domain as in a usual numerical scheme via

$$x_i = i\Delta x, \quad u_i^{(j)} := u(x_i, t_j) \quad \text{with} \quad \Delta x = \frac{b-a}{n+1},\ i = 0,1,\dots,n-1$$

and

$$t_j = j\Delta t, \quad f_i^{(j)} := f(x_i, t_j) \quad \text{with} \quad \Delta t = \frac{T}{m},\ j = 0,1,\dots,m.$$

Using the implicit Euler method in time and the shifted Grünwald method for space, we discretize the equation and obtain the following system.

$$\Big(\underbrace{\nu_{m,n}I_n + d_+T_n[g_\alpha(x)] + d_-T_n[g_\alpha(-x)]}_{T_n[f]}\Big)\mathbf{u}_n^{(j)} = \underbrace{\nu_{m,n}\mathbf{u}_n^{(j-1)} + \Delta x^\alpha \mathbf{f}_n^{(j)}}_{\mathbf{b}_n}, \tag{1.2}$$

where $\nu_{m,n} = \frac{\Delta x^\alpha}{\Delta t}$, $\mathbf{u}_n^{(j)} = [u_1^{(j)}, \ldots, u_n^{(j)}]^T$, $\mathbf{f}_n^{(j)} = [f_1^{(j)}, \ldots, f_n^{(j)}]^T$, and $g_\alpha(x) = -e^{\mathbf{i}x}(1-e^{\mathbf{i}x})^\alpha$.

The nonsymmetric $T_n[f]$, according to [15, Corollary 1], is generated by the following complex-valued function

$$f(x) = \nu_{m,n} + d_+g_\alpha(x) + d_-g_\alpha(-x).$$

Hence, the original problem of solving the FDE (1.1) becomes a classical one - solving a nonsymmetric Toeplitz linear system (1.2). Along this direction of research, most existing work such as [18,27] concerns the use of circulant preconditioners with the conjugate gradient normal residual (PCGRN) method or the preconditioned generalized minimal residual (GMRES) method for $T_n[f]$.

While such a circulant preconditioning approach has been observed effective for well-conditioned nonsymmetric Toeplitz systems, e.g. [27], its performance for the ill-conditioned ones is usually unsatisfactory as will be explained in Section 3.

Motivated by overcoming such a limitation of circulant preconditioners, we propose in this work two simple yet effective band-Toeplitz type preconditioners for a nonsymmetric ill-conditioned Toeplitz system $T_n\mathbf{x}_n = \mathbf{b}_n$. Namely, we first symmetrize the system by using Y_n without considering its normal equations system. Based on the generating function of T_n, we then construct a symmetric positive definite (SPD) band-Toeplitz preconditioner $T_n[p]$ with bandwidth $2l-1$ for the symmetrized system $Y_nT_n\mathbf{x}_n = Y_n\mathbf{b}_n$. Furthermore, we prove that the eigenvalues of the preconditioned matrix $T_n[p]^{-1}Y_nT_n$ are contained in two disjoint intervals around ± 1 with a number of outliers under certain conditions. To further speed up convergence, we also proposed a SPD band-Toeplitz plus circulant preconditioner P_n that can render clustered spectra around ± 1. As a result, rapid convergence of MINRES can be achieved. Finally, a concise analysis regarding the multilevel (block) setting is also provided, by showing the potential and the limitations of our approach.

We emphasize that by symmetrizing T_n one has rigorous theory to guide the design of preconditioners for Y_nT_n, while preconditioners for nonsymmetric problems with GMRES are often constructed based on heuristics [46]. For the sake of completeness, in the preconditioned non-Hermitian Toeplitz setting, we recall that specific rigorous tools for localizing the spectrum are studied in [25,14,16]: these mathematical tools represent a generalization of those in [7, 13,39], but they are somehow technically difficult to use in practice.

We remark that a similar approach is considered in [36], where the starting point is the similar but while we focus more specifically on preconditioning of band plus circulant type. The author [36] developed specific tools for the multilevel setting (see also the last part of Section 5).

The paper is organized as follows. In Section 2, we review the preliminary results on Toeplitz matrices. The disadvantage of circulant preconditioners for ill-conditioned Toeplitz systems is then discussed in Section 3. We provide in Section 4 guidelines on designing effective band-Toeplitz preconditioners. In Section 5, we present our main results, including a short discussion on the multilevel (block) setting. Numerical tests are given in Section 6 to support the our proposed preconditioners.

2 Preliminaries on Toeplitz matrices

Developing fast solvers for Toeplitz systems has been a research focus in the literature for their crucial applications such as numerical partial and ordinary equations, image and signal processing, etc. We refer to [8, 29, 4] for more applications.

In the context of iterative methods for Toeplitz systems, the given Toeplitz matrix $T_n[f]$ is often associated with a generating function f and different preconditioners have been developed for $T_n[f]$ based on the characteristics of f. When f is positive and continuous, the corresponding $T_n[f]$ is Hermitian positive definite (HPD) and well-conditioned. In this case, Krylov subspace methods like the conjugate gradient (CG) method with circulant preconditioners [8], τ preconditioners [2], or Hartley preconditioners [3] can achieve superlinear convergence. In the case where f is nonnegative (i.e. f has zeros), the corresponding HPD $T_n[f]$ becomes ill-conditioned [7]. While most circulant preconditioners do not work well, best circulant preconditioners proposed in [11, 9] for instance were shown to be effective in this case.

Instead of using circulant preconditioners, R. Chan [7] and Di Benedetto, et al. [13] proposed band-Toeplitz preconditioners for HPD ill-conditioned Toeplitz systems and showed their effectiveness. These band-Toeplitz preconditioners are constructed by using certain polynomial that matches the zeros of f, as reported in [7, 13]. In [10], R. Chan and Tang further improved this strategy by proposing band-Toeplitz preconditioners generated by p, where p is a trigonometric polynomial that does not only match the zeros of f but also minimizes the relative error $\|(f-p)/p\|_\infty$. Based on such an idea, Serra [40], and Noutsos and Vassalos [33, 34] later adopted different techniques to construct p for designing effective band-Toeplitz preconditioners. Other related approaches such as Toeplitz preconditioners based on the inverse approximation of Toeplitz matrices [6] have also been developed. For two-level Toeplitz systems, band two-level Toeplitz preconditioners have been used first by Serra [39] and then by Ng [30].

Throughout this work, we assume that the given Toeplitz matrix $T_n[f] \in \mathbb{C}^{n\times n}$ is associated with a Lebesgue integrable function f via its Fourier series

$$f(x) \sim \sum_{k=-\infty}^{\infty} a_k e^{\mathbf{i}kx}$$

defined on $[-\pi, \pi]$. Thus,

$$T_n[f] = \begin{bmatrix} a_0 & a_{-1} & \cdots & a_{-n+2} & a_{-n+1} \\ a_1 & a_0 & a_{-1} & & a_{-n+2} \\ \vdots & a_1 & a_0 & \ddots & \vdots \\ a_{n-2} & & \ddots & \ddots & a_{-1} \\ a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \end{bmatrix},$$

where

$$a_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-\mathbf{i}kx}\, dx, \qquad k = 0, \pm 1, \pm 2, \dots$$

are the Fourier coefficients of f. The function f is called the *generating function/spectral symbol* of $T_n[f]$. If f is complex-valued, then $T_n[f]$ is in general non-Hermitian for all n. If f is real-valued, then $T_n[f]$ is Hermitian for all n. If f is real-valued and nonnegative but not identically zero almost everywhere, then $T_n[f]$ is Hermitian positive definite for all n. If f is real-valued and even, $T_n[f]$ is real symmetric for all n [29].

We begin by introducing the following notation and definitions concerning the asymptotic singular value and spectral distribution of $T_n[f]$ associated with f.

Let $\mathcal{C}_c(\mathbb{C})$ (or $\mathcal{C}_c(\mathbb{R})$) be the space of complex-valued continuous functions defined on $\mathbb{C}$ (or $\mathbb{R}$) with bounded support and let ϕ be a functional, i.e. any function defined on some vector space which takes values in $\mathbb{C}$. Also, if $g : D \subset \mathbb{R}^k \to \mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$) is a measurable function defined on a set D with $0 < \mu_k(D) < \infty$, the functional η_g is denoted such that

$$\eta_g : \mathcal{C}_c(\mathbb{K}) \to \mathbb{C} \qquad \text{and} \qquad \eta_g(F) = \frac{1}{\mu_k(D)}\int_D F\big(g(\mathbf{x})\big)\, d\mathbf{x}.$$

Definition 2.1 [20, Definition 3.1](Singular value and eigenvalue distribution of a matrix-sequence) Let $\{A_n\}_n$ be a matrix-sequence.

1. We say that $\{A_n\}_n$ has an asymptotic singular value distribution described by a functional $\eta : C_c(\mathbb{R}) \to \mathbb{C}$, and we write $\{A_n\}_n \sim_\sigma \eta$, if

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} F(\sigma_j(A_n)) = \eta(F), \quad \forall F \in C_c(\mathbb{R}).$$

   If $\eta = \eta_{|f|}$ for some measurable $f : D \subset \mathbb{R}^k \to \mathbb{C}$ defined on a set D with $0 < \mu_k(D) < \infty$, we say that $\{A_n\}_n$ has an asymptotic singular value distribution described by f and we write $\{A_n\}_n \sim_\sigma f$.

2. We say that $\{A_n\}_n$ has an asymptotic eigenvalue (or spectral) distribution described by a functional $\eta : C_c(\mathbb{R}) \to \mathbb{C}$, and we write $\{A_n\}_n \sim_\lambda \eta$, if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} F(\lambda_j(A_n)) = \eta(F), \quad \forall F \in C_c(\mathbb{C}).$$

If $\eta = \eta_f$ for some measurable $f : D \subset \mathbb{R}^k \to \mathbb{C}$ defined on a set D with $0 < \mu_k(D) < \infty$, we say that $\{A_n\}_n$ has an asymptotic eigenvalue (or spectral) distribution described by f and we write $\{A_n\}_n \sim_\lambda f$.

Since the classical Szegő theorem was established in [38], this crucial result on the spectral distribution of Toeplitz matrices has undergone many extensions and generalizations over the years. Namely, the theorem states that the eigenvalues of the Toeplitz matrix $T_n[f]$ generated by a real-valued $f \in L^\infty([-\pi,\pi])$ are asymptotically distributed as f. Considering the same class of f, the Avram-Parter theorem [1,35] shows that the singular values of $T_n[f]$ are asymptotically distributed as $|f|$. Tyrtyshnikov later generalized the distribution result to the p-level Toeplitz matrices generated by (complex-valued) $f \in L^2([-\pi,\pi]^p)$ in [45] and later extended it to $f \in L^1([-\pi,\pi]^p)$ in [44].

The generalized Szegő theorem that describes the singular value and spectral distribution of Toeplitz sequences is given as follows:

Theorem 2.1 *(Generalized Szegő theorem)[38] Suppose $f \in L^1([-\pi,\pi])$. Let $T_n[f]$ be the Toeplitz matrix generated by f. Then*

$$\{T_n[f]\}_n \sim_\sigma f.$$

If moreover f is real-valued, then

$$\{T_n[f]\}_n \sim_\lambda f.$$

Moreover, the spectrum of Hermitian Toeplitz matrices is known to be bounded via the following theorem by R. Chan [7, Lemma 1] and by Di Benedetto et al. [13, Theorem 3.1] in which the strict inequalities are also proved.

Theorem 2.2 *[7,13] Suppose $f \in L^1([-\pi,\pi])$ is real-valued. Let m_f and M_f be the essential infimum and the essential supremum of f on $[-\pi,\pi]$, respectively, and let $T_n[f] \in \mathbb{C}^{n\times n}$ be the Toeplitz matrix generated by f. If $m_f < M_f$, then for all $n > 0$*

$$m_f < \lambda_k(T_n[f]) < M_f,$$

where λ_k is the k-th eigenvalue of $T_n[f]$ arranged in nondecreasing order. Moreover, if $m_f \geq 0$, then $T_n[f]$ is Hermitian positive definite for all n.

Throughout this work, we assume that the generating function $f \in \mathcal{C}[-\pi, \pi]$ is given with real Fourier coefficients and is periodically extended to the real line, where $\mathcal{C}[-\pi, \pi]$ is the Banach space of continuous complex-valued functions defined on $[-\pi, \pi]$. Thus, the corresponding Toeplitz matrix $T_n[f]$ is nonsymmetric in general for all n.

For (real) nonsymmetric Toeplitz matrices $T_n[f]$, while much work in the literature has been focused on using CG for their normal equations system, the authors in [37] indicated that the normalization is in fact unnecessary. Namely, one can first premultiply $T_n[f]$ by the flip/anti-identity matrix Y_n and then handle the symmetrized matrix $Y_nT_n[f]$ (i.e. a Hankel matrix) instead. One can then use MINRES in cooperation with a suitable preconditioner for $Y_nT_n[f]$ with guaranteed convergence depending only on eigenvalues, without needing to normalize the original matrix T_n. The same techniques were later shown by Hon to be applicable for functions of Toeplitz matrices [24]. We refer to [22, 21] for more on preconditioning for Toeplitz analytic function systems.

We premultiply $T_n[f]$ by using the flip matrix $Y_n \in \mathbb{R}^{n\times n}$, defined as

$$Y_n = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix},$$

and obtain the symmetric matrix

$$Y_nT_n[f] = \begin{bmatrix} a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ a_{n-2} & & \iddots & \iddots & a_{-1} \\ \vdots & a_1 & a_0 & \iddots & \vdots \\ a_1 & a_0 & a_{-1} & & a_{-n+2} \\ a_0 & a_{-1} & \cdots & a_{-n+2} & a_{-n+1} \end{bmatrix}.$$

The asymptotic spectral distribution of $Y_nT_n[f]$ that allows precise convergence analysis was first discovered in [23] and proven precisely in [19, 28] that the eigenvalues of $Y_nT_n[f]$ are distributed as $\pm|f|$ provided that $f \in L^1([-\pi, \pi])$.

Given $D \subset \mathbb{R}^k$ with $0 < \mu_k(D) < \infty$, we define $\tilde{D}$ as $D \bigcup D_r$, where $r \in \mathbb{R}^k$ and $D_r = r + D$, with the constraint that D and D_r have non-intersecting interior part, that is $D^\circ \bigcap D_r^\circ = \emptyset$. In this way $\mu_k(\tilde{D}) = 2\mu_k(D)$. Given any g defined over D, we define ψ_g over $\tilde{D}$ in the following way

$$\psi_g(x) = \begin{cases} g(x), & x \in D, \\ -g(x-r), & x \in D_r,\ x \notin D. \end{cases}$$

Theorem 2.3 *[19, Theorem 3.2] Suppose $f \in L^1([-\pi, \pi])$ with real Fourier coefficients. Let $T_n[f] \in \mathbb{R}^{n\times n}$ be the Toeplitz matrix generated by f and let $Y_n \in \mathbb{R}^{n\times n}$ be the anti-identity matrix. Then*

$$\{Y_nT_n[f]\}_n \sim_\lambda \psi_{|f|}$$

over the domain $\tilde{D}$ with $D = [0, 2\pi]$ and $r = -2\pi$.

Hence, we expect from Theorem 2.3 that $Y_n T_n[f]$ becomes ill-conditioned when $|f|$ has zeros. Also, $Y_n T_n[f]$ is in general symmetric indefinite since roughly half of its eigenvalues are negative/positive except a number of eigenvalue outliers [23], which explains and justifies the use of MINRES for the symmetrized matrix. Note that the spectral distribution of $\{Y_n T_n[f]\}_n$ is significantly different to that of $\{T_n[f]\}_n$ provided in Theorem 2.1 (the generalized Szegő theorem), even though their singular value distributions are equivalent.

The following indicates that the singular values of $Y_n T_n[f]$ (or $T_n[f]$) are essentially contained in $\left[\mathbf{essinf}_{[-\pi,\pi]}|f|, \mathbf{esssup}_{[-\pi,\pi]}|f|\right]$.

Definition 2.2 [45] Let $S \subseteq \mathbb{R}$ and let $\{A_n\}_n$ be a sequence of n-by-n complex matrices. We say that S is a *general cluster* for the singular values of $\{A_n\}_n$ if for any $\epsilon > 0$

$$\lim_{n\to\infty} \frac{\Gamma_\epsilon(n)}{n} = 0.$$

Also, S is a *proper cluster* if for any $\epsilon > 0$ there exists a constant $c(\epsilon)$ independent of n so that

$$\Gamma_\epsilon(n) \le c(\epsilon),$$

where $\Gamma_\epsilon(n)$ denotes the number of singular values of A_n whose distance from S is greater than ϵ.

If $S = \{0\}$, we say $\{A_n\}_n$ is *clustered around zero.*

As shown in [5, Corollary 4.1], the set $\left[\mathbf{essinf}_{[-\pi,\pi]}|f|, \mathbf{esssup}_{[-\pi,\pi]}|f|\right]$ is a general cluster for the singular values of $\{T_n[f]\}_n$ if $f \in L^1([-\pi,\pi])$. For instance, if f is continuous, $\left[\mathbf{essinf}_{[-\pi,\pi]}|f|, \mathbf{esssup}_{[-\pi,\pi]}|f|\right]$ is a proper cluster by [5, Theorem 4.6]. The following theorem indicates that there is no singular value of $T_n[f]$ that can be larger than $\mathbf{esssup}_{[-\pi,\pi]}|f|$.

Theorem 2.4 *[5, Theorem 2.1] Suppose $f \in L^1([-\pi,\pi])$. Let $T_n[f] \in \mathbb{C}^{n\times n}$ be the Toeplitz matrix generated by f. Then,*

$$\sigma_{\max}(T_n) \le \mathbf{esssup}_{[-\pi,\pi]}|f| \qquad \forall n \in \mathbb{N}.$$

Although $\left[\mathbf{essinf}_{[-\pi,\pi]}|f|, \mathbf{esssup}_{[-\pi,\pi]}|f|\right]$ is a singular value cluster for $\{T_n[f]\}_n$, it must however note that its singular value can in fact be smaller than $\mathbf{essinf}_{[-\pi,\pi]}|f|$. The following classical example illustrates the point: considering the Toeplitz matrix

$$T_n[e^{\mathbf{i}x}] = \begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n\times n},$$

we can see that its smallest singular value is 0 but $\mathbf{essinf}_{[-\pi,\pi]}|e^{\mathbf{i}x}| = 1$.

Nevertheless, as indicated in [5], the number of the singular values of $T_n[f]$ that are smaller than $\mathbf{essinf}_{[-\pi,\pi]}|f|$ cannot grow linearly with n. It was also shown in the same paper that if f is weakly sectorial and not pathological, the minimal singular value of the corresponding $T_n[f]$ is bounded below. Thus, the generated Toeplitz matrix in this case cannot be arbitrarily ill-conditioned (see [5, Section 3] for detail).

Unless stated otherwise, we therefore in this work consider $T_n[f]$ as ill-conditioned if $\mathbf{essinf}_{[-\pi,\pi]}|f| = 0$ regardless of the location of outliers which might result in arbitrarily ill-conditioning. Also, we require throughout that the continuous generating function f is given such that $|f|$ has zeros of even order in order to construct an effective band-Toeplitz preconditioner for $Y_nT_n[f]$, which will be discussed in detail in Section 4.

3 Limitation of circulant matrices as preconditioners

In this section, we briefly discuss the use of circulant matrices as preconditioners for Toeplitz systems and their limitation in this regard.

We first provide the following definitions.

Definition 3.1 [37] For any circulant matrix $C_n \in \mathbb{C}^{n\times n}$, the *absolute value circulant matrix* $|C_n|$ of C_n is defined by

$$|C_n| = (C_n^*C_n)^{1/2} = (C_nC_n^*)^{1/2} = F_n^*|\Lambda_n|F_n,$$

where $F_n \in \mathbb{C}^{n\times n}$ is the Fourier matrix and $|\Lambda_n|$ is the diagonal matrix in the eigendecomposition of C_n with all entries replaced by their magnitude.

After symmetrizing a real Toeplitz matrix T_n by using Y_n, Pestana and Wathen [37] proposed an absolute value circulant preconditioner $|C_n|$ for the symmetrized matrix Y_nT_n with C_n derived from T_n in a standard way. The authors then concluded in [37, Proposition 4.1] that $|C_n|^{-1}Y_nT_n$ has clustered spectra around ±1 by showing the following: for all $\epsilon > 0$ there exist integers M and N such that for all $n > N$

$$|C_n|^{-1}Y_nT_n = Q_n + \tilde{R}_n + \tilde{E}_n,$$

where Q_n is involutory (i.e. symmetric and orthogonal), rank $\tilde{R}_n \leq 2M$, and $\|\tilde{E}_n\|_2 \leq \epsilon$. Indeed, the clustered spectrum result follows as precisely shown in [19, Theorem 3.2].

However, such a matrix decomposition is in fact insufficient to guarantee rapid convergence when T_n is ill-conditioned, namely when $\mathbf{essinf}_{[-\pi,\pi]}|f| = 0$ (i.e. when $|f|$ has zeros as discussed in Section 2). Even though the eigenvalues of $|C_n|^{-1}Y_nT_n$ are clustered around ±1, the smallest one is not necessarily uniformly bounded away from zero with respect to n.

In other words, even though we have a way to deal with nonsymmetric Toeplitz matrix T_n (via the flip matrix Y_n) with guaranteed convergence depending only on eigenvalues, the quality of (absolute value) circulant preconditioners remains unsatisfactory for ill-conditioned T_n. Due to such a drawback of circulant preconditioners, we devote to constructing band-Toeplitz preconditioners for Y_nT_n in this work.

4 Band-Toeplitz preconditioners

In this section, we provide some guidelines on designing effective band-Toeplitz preconditioners for $Y_nT_n[f]$.

As mentioned in Section 2, the asymptotic spectral distribution of $Y_nT_n[f]$ is essentially $\pm|f|$. Accordingly, we construct a band-Toeplitz preconditioner for $Y_nT_n[f]$ by finding a trigonometric polynomial of order $l \geq k$ that approximates $|f|$ and matches its zeros on $[-\pi, \pi]$. Note that we adopt in here the classical proposals [7,40] as examples.

Let p be the generating function of our proposed band-Toeplitz preconditioner and let z_k be the polynomial of minimal degree k that contains all the zeros of $|f|$. We adopt in this work the following approximation approach as an example: to find p such that

$$p = z_k p_{l-k},$$

where p_{l-k} is the best approximation of degree $l-k$ to $|f|/z_k$, i.e.

$$\left\| \frac{|f|}{z_k} - p_{l-k} \right\|_\infty = \min_{g \in \mathbb{P}_{l-k}} \left\| \frac{|f|}{z_k} - g \right\|_\infty .$$

Suppose $|f|$ has zeros $x_1, \ldots, x_m$ of order $2l_1, \ldots, 2l_m$, respectively. As proposed in [7], we can choose the following trigonometric polynomial z_k to match the zeros of $|f|$

$$z_k(x) = \prod_{i=1}^{m} \big(2 - 2\cos(x - x_i)\big)^{l_i}, \qquad k = \sum_{i=1}^{m} l_i. \tag{4.1}$$

Note that p_{l-k} can be computed by the standard Remez algorithm since $|f|/z_k$ is a positive, continuous function on $[-\pi, \pi]$.

Let $\mathbb{P}_l$ be the set of trigonometric polynomial of degree at most l. The function $|f|$ has the following property.

Theorem 4.1 *[40, Theorem 4.1] Let $|f|$ be a nonnegative continuous function defined on $[-\pi, \pi]$ with zeros of even order. Then, there exists a nonnegative trigonometric polynomial z_k of minimal degree k such that*

$$0 < r_k < \frac{|f|}{z_k} < R_k < \infty.$$

Alternatively, one can for example choose p to be the following approximation to $|f|$:

1. The best approximation proposed in [10], i.e. the trigonometric polynomial p such that
$$\left\| \frac{|f| - p}{p} \right\|_\infty = \min_{p^* \in \mathbb{P}_l} \left\| \frac{|f| - p^*}{p^*} \right\|_\infty .$$

   Remark 4.1 Since $|f|$ has zeros, the standard Remez algorithm cannot be straightforwardly applied as mentioned in [10].

2. The trigonometric polynomial originally proposed in [7], i.e. $p = z_k$.
3. The trigonometric polynomial proposed in [40] such that $p = z_k p_{l-k}$, where p_{l-k} is the trigonometric polynomial interpolating $|f|/z_k$ at the $l-k+1$ zeros of the $(l-k+1)$-th Chebyshev polynomial of the first kind.
4. The technique proposed in [33], namely constructing a rational approximation of $|f|/z_k$.
5. The Toeplitz plus circulant technique proposed in [34].

Our proposed preconditioning here is not restricted to the abovementioned strategies. Based on the characteristic of $|f|$, one can accordingly design an effective preconditioner for $Y_n T_n[f]$.

5 Main result

In this section, we present the following main theorem accounts for the effectiveness of our proposed band-Toeplitz preconditioners for $Y_n T_n[f]$.

5.1 Band-Toeplitz preconditioners

Theorem 5.1 *Suppose $f \in \mathcal{C}[-\pi, \pi]$ with real Fourier coefficients and $|f|$ only has zeros of even order. Let $T_n[f] \in \mathbb{R}^{n\times n}$ be the Toeplitz matrix generated by f, let $Y_n \in \mathbb{R}^{n\times n}$ be the anti-identity matrix, and let $T_n[p] \in \mathbb{R}^{n\times n}$ be the Toeplitz matrix generated by the trigonometric polynomial p of degree l, i.e. $p = z_k p_{l-k}$ where z_k is defined as in (4.1) and p_{l-k} is the best approximation of degree $l-k$ to $|f|/z_k$ such that*

$$\left\| \frac{|f|}{z_k} - p_{l-k} \right\|_\infty =: h_{l-k}.$$

Then,

1. *$T_n[p]$ is symmetric positive definite,*
2. *$T_n[p]^{-1} Y_n T_n[f]$ has $n/2+\mathcal{O}(1)$ eigenvalues contained in $[-\beta, -\alpha]$ and $n/2+\mathcal{O}(1)$ eigenvalues contained in $[\alpha, \beta]$, where*

$$\alpha = 1 - \frac{h_{l-k}}{r_k + h_{l-k}} \qquad \text{and} \qquad \beta = 1 + \frac{h_{l-k}}{r_k - h_{l-k}}$$

with $r_k = \min_{[-\pi,\pi]} (|f|/z_k)$.

Proof We divide this proof into the following parts.

Proof of Part I

By the assumption that $T_n[f]$ is real and the fact it is uniquely defined by its symmetric and skew-symmetric parts, we have equivalently required that its generating function has the form $f = f_R + \mathbf{i} f_I$, where f_R is a real-valued even function and f_I is a real-valued odd one. In other words, $|f|$ is an even function which means p is also even by our construction. Also, as p is nonnegative, $T_n[p]$ is symmetric positive definite for all n by Theorem 2.2.

Proof of Part II

We first examine $T_n[p]^{-1}$. Since $T_n[p]$ is symmetric, we have $T_n[p] = Y_n T_n[p] Y_n$, which implies

$$T_n[p]^{-1} = Y_n T_n[p]^{-1} Y_n. \tag{5.1}$$

As p is a trigonometric polynomial of degree l, using a standard result, we have

$$T_n[f] = T_n[p] T_n[f/p] + R_n, \tag{5.2}$$

where rank $R_n = 2l$ (e.g. see [29, Theorem 6.5]),

By (5.1) and (5.2), we write

$$\begin{aligned} T_n[p]^{-1/2} Y_n T_n[f] T_n[p]^{-1/2} &= T_n[p]^{1/2} T_n[p]^{-1} Y_n T_n[f] T_n[p]^{-1/2} \\ &= T_n[p]^{1/2} Y_n T_n[p]^{-1} T_n[f] T_n[p]^{-1/2} \\ &= T_n[p]^{1/2} Y_n T_n[p]^{-1} \Big(T_n[p] T_n[f/p] + R_n \Big) T_n[p]^{-1/2} \\ &= T_n[p]^{1/2} Y_n T_n[f/p] T_n[p]^{-1/2} + R_n^{(1)}, \end{aligned}$$

where rank $R_n^{(1)} \leq 2l$ independent of n. Hence, $T_n[p]^{-1/2} Y_n T_n[f] T_n[p]^{-1/2}$ (being similar to $T_n[p]^{-1} Y_n T_n[f]$) and $Y_n T_n[f/p]$ have the same asymptotic singular distribution.

We now investigate the singular values of $T_n[p]^{-1} Y_n T_n[f]$ by those of $T_n[f/p]$. Recall that p is a trigonometric polynomial such that $p = z_k p_{l-k}$, where p_{l-k} is best approximation of degree $l-k$ to $|f|/z_k$, i.e.

$$\Big\| \underbrace{\frac{|f|}{z_k}}_{\hat{f}} - p_{l-k} \Big\|_\infty = h_{l-k}. \tag{5.3}$$

From (5.3), we have

$$-h_{l-k} \leq \hat{f} - p_{l-k} \leq h_{l-k}. \tag{5.4}$$

Dividing (5.4) by $\hat{f} > 0$ and then rearranging the terms gives

$$\frac{\hat{f}}{\hat{f} + h_{l-k}} \leq \frac{\hat{f}}{p_{l-k}} \leq \frac{\hat{f}}{\hat{f} - h_{l-k}}.$$

Since $\hat{f} = |f|/z_k$ is positive and continuous on $[-\pi, \pi]$, $\min_{[-\pi,\pi]} \hat{f} =: r_k > 0$ exists by Theorem 4.1 and we have

$$\begin{aligned} 1 - \frac{h_{l-k}}{\hat{f} + h_{l-k}} &\leq \frac{\hat{f}}{p_{l-k}} \leq 1 + \frac{h_{l-k}}{\hat{f} - h_{l-k}} \\ 1 - \frac{h_{l-k}}{r_k + h_{l-k}} &\leq \underbrace{\frac{\hat{f}}{p_{l-k}}}_{\frac{|f|}{p}} \leq 1 + \frac{h_{l-k}}{r_k - h_{l-k}}. \end{aligned}$$

Therefore, we conclude all singular values of $T_n[f/p]$ are essentially contained in the interval $\big[\alpha, \beta\big]$, where

$$\alpha = 1 - \frac{h_{l-k}}{r_k + h_{l-k}} \quad \text{and} \quad \beta = 1 + \frac{h_{l-k}}{r_k - h_{l-k}}.$$

As $T_n[p]^{-1/2} Y_n T_n[f] T_n[p]^{-1/2}$ is symmetric, its eigenvalues are contained in the interval $[-\beta, -\alpha] \cup [\alpha, \beta]$ except a number of outliers for large enough n. Since the disjoint interval is a proper cluster for continuous f/p and rank R_n is independent of n, we further deduce that the number of eigenvalue outliers of $T_n[p]^{-1/2} Y_n T_n[f] T_n[p]^{-1/2}$ is fixed (with respect to n).

By Sylvester's law of inertia, both $T_n[p]^{-1} Y_n T_n[f]$ and $Y_n T_n[f]$ share the same numbers of positive, negative, and zero eigenvalues. Consequently, as $\{Y_n T_n[f]\}_n \sim_\lambda \psi_{|f|}$ by Theorem 2.3, we know that $T_n[p]^{-1} Y_n T_n[f]$ has $n/2 + \mathcal{O}(1)$ eigenvalues contained in $[-\beta, -\alpha]$ and $n/2 + \mathcal{O}(1)$ eigenvalues contained in $[\alpha, \beta]$.

As a consequence of Theorem 5.1, we have the following scenario:

1. If all eigenvalues of $T_n[p]^{-1} Y_n T_n[f]$ are contained in the intervals $[-\beta, -\alpha] \cup [\alpha, \beta]$ with no outliers, superlinear convergence is achieved by a well-known classical result on MINRES convergence for example in [17]. Namely, the k-th residual of $\mathbf{r}_n^{(k)}$ satisfies $\frac{\|\mathbf{r}_n^{(k)}\|_2}{\|\mathbf{r}_n^{(0)}\|_2} \leq 2\left(\frac{\beta/\alpha - 1}{\beta/\alpha + 1}\right)^{[k/2]}$. It is clear in this case that the number of iterations MINRES required to converge is independent of n.
2. If there are outliers and they are uniformly bounded away from zero, superlinear convergence is still obtainable by [48, Theorem 3.2].

Using a similar argument, we can further weaken the assumption on f and generalize Theorem 5.1 to the following result.

Theorem 5.2 *Suppose $f \in L^\infty([-\pi, \pi])$ with real Fourier coefficients and $|f|$ is essentially positive. Let $T_n[f] \in \mathbb{R}^{n \times n}$ be the Toeplitz matrix generated by f, let $Y_n \in \mathbb{R}^{n \times n}$ be the anti-identity matrix, and let $T_n[|f|] \in \mathbb{R}^{n \times n}$ be the Toeplitz matrix generated by $|f|$. Then,*

1. *$T_n[|f|]$ is symmetric positive definite and*
2. *the eigenvalues of $T_n[|f|]^{-1} Y_n T_n[f]$ are clustered around ± 1, and*
3. *$T_n[|f|]^{-1} Y_n T_n[f]$ has $n/2 + o(n)$ eigenvalues clustered around 1 and $n/2 + o(n)$ eigenvalues clustered around -1.*

Proof This proof is similar to that of Theorem 5.1 and yet the involved technicalities are somewhat different. For completeness, we provide a proof in the following.

Proof of Part I

As $|f|$ is an even, essentially positive function (see the proof of Part I in Theorem 5.1), $T_n[|f|]$ is SPD.

Proof of Part II

As $T_n[|f|]$ is SPD, we have

$$T_n[|f|]^{-1} = Y_n T_n[|f|]^{-1} Y_n. \tag{5.5}$$

Since $\|T_n[|f|]^{-1}T_n[f] - T_n[f/|f|]\|_* = o(n)$ by [14, Proposition 4], where $\|\cdot\|_*$ is the trace norm and $o(\cdot)$ is the standard small-o notation. Note that a similar argument is also used in the proof of [36, Theorem 3.7]. Further with (5.5), we write

$$\begin{aligned} T_n[|f|]^{-1/2} Y_n T_n[f] T_n[|f|]^{-1/2} &= T_n[|f|]^{1/2} Y_n T_n[|f|]^{-1} T_n[f] T_n[|f|]^{-1/2} \\ &= T_n[|f|]^{1/2} Y_n T_n[f/|f|] T_n[|f|]^{-1/2} + \Delta_n, \end{aligned}$$

where $\{\Delta_n\}_n \sim_{\lambda,\sigma} 0$.

Thus, we know $T_n[|f|]^{-1}Y_nT_n[f]$ (being similar to $T_n[|f|]^{-1/2}Y_nT_n[f]T_n[|f|]^{-1/2}$) and $Y_nT_n[f/|f|]$ have the same asymptotic singular distribution.

As the singular values of $Y_nT_n[f/|f|]$ are clustered around one and the low rank matrix hidden in Δ_n is of $o(n)$ rank, the eigenvalues of $T_n[|f|]^{-1}Y_nT_n[f]$ are clustered around ± 1. By Sylvester's law of inertia and Theorem 2.3, we can further show that $T_n[|f|]^{-1}Y_nT_n[f]$ possesses $n/2 + o(n)$ eigenvalues clustered around 1 and $n/2 + o(n)$ eigenvalues clustered around -1.

In Theorem 5.2, superlinear convergence may not be obtainable in general due to the increasing eigenvalue outliers. However, the outliers cannot grow linearly with n. Provided that there are no eigenvalue outliers, optimal convergence for $T_n[|f|]^{-1}Y_nT_n[f]$ in two iterations can be achieved.

It must however note that $T_n[|f|]$ as preconditioner is unrealistic - the inversion of $T_n[|f|]$ is computationally expensive and hence the resulting complexity will be high in general. Note that in the analogy of the SPD unilevel Toeplitz case, $T_n[|f|] = T_n[f]$ is in fact the Toeplitz matrix to be inverted itself. Hence, Theorem 5.2 can be regarded as a guideline for designing good preconditioners that approximates $T_n[|f|]$ but on its own is merely of theoretical interest.

We end this subsection by giving the following result on another band-Toeplitz preconditioner $T_n[z_k]$ (i.e. we simply pick $p = z_k$), which is easy to construct and does not involve any approximation process. This preconditioner corresponds to the strategy 2 proposed in Section 4.

Theorem 5.3 *Suppose $f \in \mathcal{C}[-\pi,\pi]$ with real Fourier coefficients and $|f|$ only has zeros of even order. Let $T_n[f] \in \mathbb{R}^{n\times n}$ be the Toeplitz matrix generated by f, let $Y_n \in \mathbb{R}^{n\times n}$ be the anti-identity matrix, and let $T_n[z_k] \in \mathbb{R}^{n\times n}$ be the Toeplitz matrix generated by z_k, where z_k is defined as in (4.1). Then,*

1. *$T_n[z_k]$ is symmetric positive definite,*
2. *$T_n[z_k]^{-1}Y_nT_n[f]$ has $n/2 + \mathcal{O}(1)$ eigenvalues contained in $[-\beta,-\alpha]$ and $n/2 + \mathcal{O}(1)$ eigenvalues contained in $[\alpha,\beta]$, where*

$$\alpha = \min_{[-\pi,\pi]} (|f|/z_k) \qquad \text{and} \qquad \beta = \max_{[-\pi,\pi]} (|f|/z_k).$$

To sum up, our proposed preconditioning approach can achieve overall $\mathcal{O}(n \log n)$ complexity due to the efficient computation of $T_n[p]^{-1}\mathbf{d}_n$ for any $\mathbf{d}_n$ by a classical band solver in $\mathcal{O}(l^2 n)$ operations and the rapid convergence resulting from the narrow spectrum around ± 1.

Despite the fact that the product $T_n[p]^{-1}\mathbf{d}_n$ can be efficiently computed in $\mathcal{O}(l^2 n)$ operations, we point out that l in practice might grow with the dimension n. Since for any n the trigonometric approximation p is computed to construct $T_n[p]$, l being the degree of p can be large for irregular $|f|$. Hence, one needs to calculate p by using efficient methods in order to maintain the overall $\mathcal{O}(n \log n)$ complexity. As identified by Serra in [40, Section 6], the approximation approach 4 given in Section 4 that involves Chebyshev polynomials of the first kind is an example of such efficient methods.

5.2 Band-Toeplitz plus circulant preconditioners

Because of the mentioned disadvantage of $T_n[p]$, we also propose in this subsection the following band-Toeplitz plus circulant preconditioner that does not involve any approximation process.

The idea is classical - we eliminate the zeros of $|f|$ using the band Toeplitz matrix $T_n[z_k]$ and accelerate the convergence using a circulant preconditioner. Namely, based on, we propose the SPD preconditioner $P_n = C_n[\sqrt{|f|/z_k}]T_n[z_k]C_n[\sqrt{|f|/z_k}]$. As $\sqrt{|f|/z_k}$ is positive, a usual circulant preconditioner $C_n[\sqrt{|f|/z_k}]$ derived from $T_n[\sqrt{|f|/z_k}]$ will suffice.

Theorem 5.4 *Suppose $f \in \mathcal{C}[-\pi, \pi]$ with real Fourier coefficients and $|f|$ only has zeros of even order. Let $T_n[f] \in \mathbb{R}^{n\times n}$ be the Toeplitz matrix generated by f, let $Y_n \in \mathbb{R}^{n\times n}$ be the anti-identity matrix, and let $P_n = C_n[\sqrt{|f|/z_k}]T_n[z_k]C_n[\sqrt{|f|/z_k}] \in \mathbb{R}^{n\times n}$. Then,*

1. *P_n is symmetric positive definite and*
2. *$P_n^{-1}Y_nT_n[f]$ has $n/2+\mathcal{O}(1)$ eigenvalues clustered around 1 and $n/2+\mathcal{O}(1)$ eigenvalues clustered around -1.*

Proof We divide this proof into the following parts.

Proof of Part I

Since $|f|/z_k$ is a positive, even function by assumption, $C_n[\sqrt{|f|/z_k}]$ is a SPD circulant matrix. We know $C_n[\sqrt{|f|/z_k}]^{-1}$ can be symmetrized by Y_n, namely

$$Y_nC_n[\sqrt{|f|/z_k}]^{-1} = C_n[\sqrt{|f|/z_k}]^{-1}Y_n. \tag{5.6}$$

Also, $T_n[z_k]$ a SPD Toeplitz matrix. Hence, $P_n = C_n[\sqrt{|f|/z_k}]T_n[z_k]C_n[\sqrt{|f|/z_k}]$ being congruent with $T_n[z_k]$ is also SPD for all n by Sylvester's law of inertia.

Proof of Part II

For positive, even, continuous $\sqrt{|f|/z_k}$, it is well-known that the following matrix decomposition on $C_n[\sqrt{|f|/z_k}]$ holds (see for example [29, Theorem 4.10] for a result on the optimal circulant preconditioners [12] by T. Chan):

$$C_n[\sqrt{|f|/z_k}]^{-1}T_n[\sqrt{|f|/z_k}] = I_n + R_n + E_n, \tag{5.7}$$

where $R_n \in \mathbb{R}^{n\times n}$ is of fixed rank independent of n and $E_n \in \mathbb{R}^{n\times n}$ is of small norm. We remark that $T_n[\sqrt{|f|/z_k}]$ exists and is SPD by Theorem 2.2.

Using similar arguments in the proof of Theorem 5.1 and Eqns. (5.6) and (5.7), we have

$$
\begin{aligned}
& P_n^{-1} Y_n T_n[f] \\
&= (C_n[\sqrt{|f|/z_k}] T_n[z_k] C_n[\sqrt{|f|/z_k}])^{-1} Y_n T_n[f] \\
&= C_n[\sqrt{|f|/z_k}]^{-1} T_n[z_k]^{-1} C_n[\sqrt{|f|/z_k}]^{-1} Y_n T_n[f] \\
&= C_n[\sqrt{|f|/z_k}]^{-1} T_n[\sqrt{|f|/z_k}] T_n[\sqrt{|f|/z_k}]^{-1} T_n[z_k]^{-1} C_n[\sqrt{|f|/z_k}]^{-1} Y_n T_n[f] \\
&= \Big(I_n + R_n + E_n\Big) T_n[\sqrt{|f|/z_k}]^{-1} T_n[z_k]^{-1} C_n[\sqrt{|f|/z_k}]^{-1} Y_n T_n[f] \\
&= \Big(I_n + R_n + E_n\Big) T_n[\sqrt{|f|/z_k}]^{-1} T_n[z_k]^{-1} Y_n C_n[\sqrt{|f|/z_k}]^{-1} T_n[f] \\
&= \Big(I_n + R_n + E_n\Big) T_n[\sqrt{|f|/z_k}]^{-1} T_n[z_k]^{-1} Y_n C_n[\sqrt{|f|/z_k}]^{-1} T_n[\sqrt{|f|/z_k}] T_n[\sqrt{|f|/z_k}]^{-1} T_n[f] \\
& \quad \vdots \\
&= T_n[\sqrt{|f|/z_k}]^{-1} T_n[z_k]^{-1} Y_n T_n[\sqrt{|f|/z_k}]^{-1} T_n[f] + R_n^{(1)} + E_n^{(1)} \\
&= Y_n T_n[f/|f|] + R_n^{(2)} + E_n^{(2)},
\end{aligned}
$$

where $R_n^{(2)}$ is of a small fixed rank independent of n and $E_n^{(2)}$ is of small norm.

Thus, $P_n^{-1} Y_n T_n[f]$ and $Y_n T_n[f/|f|]$ have the same asymptotic singular distribution. Since the singular values of $Y_n T_n[f/|f|]$ are clustered around 1, the eigenvalues of $P_n^{-1} Y_n T_n[f]$ (being similar to the symmetric matrix $P_n^{-1/2} Y_n T_n[f] P_n^{-1/2}$) are clustered around ± 1 with a fixed number of outliers. By Sylvester's law of inertia and Theorem 2.3, we further show that $P_n^{-1} Y_n T_n[f]$ possesses $n/2 + \mathcal{O}(1)$ eigenvalues clustered around 1 and $n/2 + \mathcal{O}(1)$ eigenvalues clustered around -1.

As a consequence of Theorem 5.4, superlinear convergence is obtainable by [48, Theorem 3.2], provided that the outliers are uniformly bounded away from zero. Also, for any vector $\mathbf{d}_n$, $P_n^{-1}\mathbf{d}_n = C_n[\sqrt{|f|/z_k}]^{-1} T_n[z_k]^{-1} C_n[\sqrt{|f|/z_k}]^{-1}\mathbf{d}_n$ can be computed in $\mathcal{O}(n \log n)$ operations using fast Fourier transforms and a band solver. Hence, the overall $\mathcal{O}(n \log n)$ complexity of the solver with P_n can be attained.

5.3 Extension to the multilevel (block) Toeplitz case

Our proposed preconditioning strategy can be naturally extended to the block Toeplitz case and to the most general multilevel (block) Toeplitz case.

We first consider the block case: 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

$$\|f\|_{L^1} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \|f(\theta)\|_* \, d\theta < \infty$$

with

$$\|f\|_* = \frac{1}{m}\sum_{j=1}^{m}\lambda_j(|f|(\theta))) \qquad |f| = (ff^*)^{1/2}.$$

In the following, we assume that $f \in L^1([-\pi,\pi],\mathbb{C}^{m\times m})$ unless mentioned otherwise. The block Toeplitz matrix $T_n[f]$ generated by f has the matrix-valued Fourier coefficients

$$A_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-\mathbf{i}k\theta}\,d\theta \in \mathbb{C}^{m\times m}, \qquad k = 0, \pm1, \pm2, \ldots.$$

If f is Hermitian, $T_n[f]$ is Hermitian by [41, Lemma 2.1] and we refer to the paper for more related properties.

In addition, the blocks $A_{(k)}$ are assumed to be Hermitian so we can make $T_n[f]$ Hermitian by premultiplying it $Y_{(n,m)} = Y_n \otimes I_m \in \mathbb{R}^{mn\times mn}$. Also, we have the following analogous result to Theorem 2.3 on the asymptotic spectral distribution of $(Y_n \otimes I_m)T_n[f]$.

Theorem 5.5 *[19, Theorem 3.4] Suppose $f \in L^1([-\pi,\pi],\mathbb{C}^{m\times m})$ has Hermitian Fourier coefficients. Let $T_n[f] \in \mathbb{C}^{mn\times mn}$ be the block Toeplitz matrix generated by f. Then*

$$\{(Y_n \otimes I_m)T_n[f]\}_n \sim_\lambda \psi_{|f|}, \qquad |f| = (ff^*)^{1/2},$$

over the domain $\tilde{D}$ with $D = [0, 2\pi]$ and $p = -2\pi$, where $\psi_{|f|}$ is defined in (2.1). That is

$$\lim_{n\to\infty}\frac{1}{mn}\sum_{j=1}^{mn}F(\lambda_j((Y_n\otimes I_m)T_n[f])) = \frac{1}{4\pi}\int_{\tilde{D}}\frac{1}{m}\sum_{j=1}^{m}F(\lambda_j(\psi_{|f|}(\theta))\,d\theta.$$

As $f \in L^1([-\pi,\pi],\mathbb{C}^{m\times m})$, the closed interval

$$\left[\mathbf{essinf}_{[-\pi,\pi]}\lambda_{\min}(|f|), \mathbf{esssup}_{[-\pi,\pi]}\lambda_{\max}(|f|)\right]$$

is a singular value cluster for the related sequence by [5, Theorem 4.6]. If smoothness assumptions are added (the continuity is sufficient), then the cluster is of proper type. Using similar arguments given in the previous section, superlinear linear convergence rate can still be obtained in the block Toeplitz case.

Interestingly enough, if we consider the multilevel (block) Toeplitz case, then everything can be generalized plainly (see also [36]). The only price to pay is a standard multilevel notation.

Let f be a Lebesgue integrable d-variate function taking values in the space $\mathbb{C}^{m\times m}$ that is $f \in L^1([-\pi,\pi]^d,\mathbb{C}^{m\times m})$. Let $\mathbf{n} := (n_1,\ldots,n_d)$ be a multi-index in $\mathbb{N}^d$ and set $\hat{n} := \prod_{i=1}^{d} n_i$.

Now let the Fourier coefficients of a given function $f \in L^1([-\pi,\pi]^d, \mathbb{C}^{m\times m})$ be defined as

$$A_{\mathbf{k}} = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} f(\theta) \mathrm{e}^{-\mathbf{i}\langle \mathbf{k},\theta\rangle} \, d\theta, \qquad \mathbf{k} = (k_1,\ldots,k_d) \in \mathbb{Z}^d, \quad \mathbf{i}^2 = -1, \tag{5.8}$$

where $\langle \mathbf{k},\theta\rangle = \sum_{t=1}^{d} k_t\theta_t$ and the integrals in (5.8) are computed componentwise. Then, the $\mathbf{n}$-th Toeplitz matrix associated with f is the matrix of order $m\hat{n}$ given by

$$T_{\mathbf{n}}[f] = \sum_{\mathbf{k}=-(\mathbf{n}-\mathbf{e})}^{\mathbf{n}-\mathbf{e}} J_{n_1}^{j_1} \otimes \cdots \otimes J_{n_d}^{j_d} \otimes A_{\mathbf{k}}, \tag{5.9}$$

with J_m^l being the matrix of size m having 1 in position (j,k) such that $j-k=l$ and zero otherwise.

A generalization of Theorem 5.5 can be also obtained and it is an interesting topic in itself. However, our focus here regards the choice of proper preconditioners and we briefly discuss the main features of our band preconditioning in the multilevel (block) setting.

When considering for example an extension of Theorem 5.1 to the case where $d=1$ and $m>1$ (unilevel block case), from the key relation (5.2) we obtain

$$T_n[f] = T_n[p]T_n[f/p] + R_n,$$

with p matrix-valued trigonometric polynomial of degree l and rank $R_n = 2lm$. Now the correction matrix R_n has a rank proportional to m, and hence we expect a similar worsening in the number of outliers in the associated preconditioning. The situation becomes less nice when $d>1$. In that case, looking again at the key relation (5.2) we obtain

$$T_{\mathbf{n}}[f] = T_{\mathbf{n}}[p]T_{\mathbf{n}}[f/p] + R_{\mathbf{n}},$$

with p matrix-valued, d-variate trigonometric polynomial. In such a context the correction matrix $R_{\mathbf{n}}$ has rank proportional to the degree of p and to m but also to $\hat{n}(\sum_{i=1}^{d} n_i^{-1})$.

In other words the number of expected outliers, under the simplifying assumption that $n_1 \sim n_2 \sim \cdots \sim n_d$, is proportional to $N^{(d-1)/d}$ where $N = m\hat{n}$ is the global size of the matrix. The latter worsening already for $d=2$ prevents superlinear convergence and the nature of this negative phenomenon is not a surprise since it has been already studied in the case of matrix algebra preconditioners for multilevel (block) Toeplitz structures (see e.g. [42,31]).

6 Numerical results

In this section, we demonstrate the effectiveness of our proposed preconditioners using MINRES. Throughout all numerical tests, we use the MATLAB R2018b built-in functions **minres** and (nonrestarted) **gmres**, and run on a dual-core, 1.4 GHz Intel i5 CPU with 4 GB RAM. The vector $\mathbf{b}_n$ is generated by **ones(n,1)**, the initial guess is the zero vector, and the stopping criterion used is $\frac{\|\mathbf{r}_n^{(j)}\|_2}{\|\mathbf{r}_n^{(0)}\|_2} < 10^{-7}$, where $\mathbf{r}_n^{(j)}$ is the residual vector after j iterations. Note that the Remez algorithm is executed by the Chebfun (V5.7.0) built-in function **trigremez**. The existing circulant preconditioners $s_n[f]$ and $c_n[f]$ for $T_n[f]$ used in the tests are Strang's circulant preconditioners [43] and optimal (T. Chan's) circulant preconditioners for $T_n[f]$ respectively (see [26] for a survey on the latter and [47] for a related fast solver based on circulant structures).

Example 6.1 We first consider the ill-conditioned Toeplitz matrix $T_n[f]$ generated by $f(x) = x^2 + \mathbf{i}x^3$. Simple calculations yield

$$|f(x)| = x^2\sqrt{1+x^2}$$

which has a zero at $x = 0$ of order 2. Thus, we pick

$$z_1(x) = 2 - 2\cos x$$

of degree $k = 1$.

Table 6.1 shows the iteration numbers MINRES needed for $Y_nT_n[f]$ with the preconditioners I_n (i.e. with no preconditioner), the existing absolute value circulant preconditioners $|s_n[f]|$ and $|c_n[f]|$, and our proposed preconditioners $T_n[p]$, $T_n[z_1]$, $P_{n,s} := s_n[\sqrt{|f|/z_1}]T_n[z_1]s_n[\sqrt{|f|/z_1}]$, and $P_{n,c} := c_n[\sqrt{|f|/z_1}]T_n[z_1]c_n[\sqrt{|f|/z_1}]$. Compared with both $|s_n[f]|$ and $|c_n[f]|$, we observe that our proposed preconditioners are more effective for speeding up the convergence rate of MINRES. As a comparison, Table 6.2 shows the GMRES iteration numbers with $s_n[f]$, $c_n[f]$, and our proposed preconditioners. It is remarked that GMRES with $T_n[z_1]$ was proposed in [25, Example 3]. Again, our preconditioners appear to have lower iteration numbers and P_n works best. In Figure 6.1, we observe the expected spectra with our preconditioners when $n = 512$. A rapid convergence rate of MINRES with our proposed preconditioners is then obtained.

Also, it is worth noticing that a GMRES method with band-Toeplitz preconditioner was developed in [32, Example 2], where the authors proved singular value clusters. In contrast, our MINRES method can attain guaranteed convergence which is based only on eigenvalues clusters.

Table 6.1 Numbers of MINRES iterations for $Y_nT_n[f]$ with $T_n[f]$ generated by $f(x) = x^2 + \mathbf{i}x^3$. Note that n/a means MATLAB fails to converge because the method stagnated.

	MINRES						
n	I_n	$\lvert s_n[f]\rvert$	$\lvert c_n[f]\rvert$	$T_n[z_1]$	$T_n[p]$	$P_{n,s}$	$P_{n,c}$
512	>1000	251	71	144	37	12	15
1024	>1000	532	100	153	37	14	15
2048	>1000	>1000	191	159	39	15	15
4096	>1000	>1000	n/a	163	39	15	15

Table 6.2 Numbers of GMRES iterations for $T_n[f]$ generated by $f(x) = x^2 + \mathbf{i}x^3$.

	GMRES						
n	I_n	$s_n[f]$	$c_n[f]$	$T_n[z_1]$	$T_n[p]$	$P_{n,s}$	$P_{n,c}$
512	512	88	31	58	24	19	19
1024	>1000	122	39	53	22	17	17
2048	>1000	168	52	46	20	15	15
4096	>1000	233	68	40	18	13	13

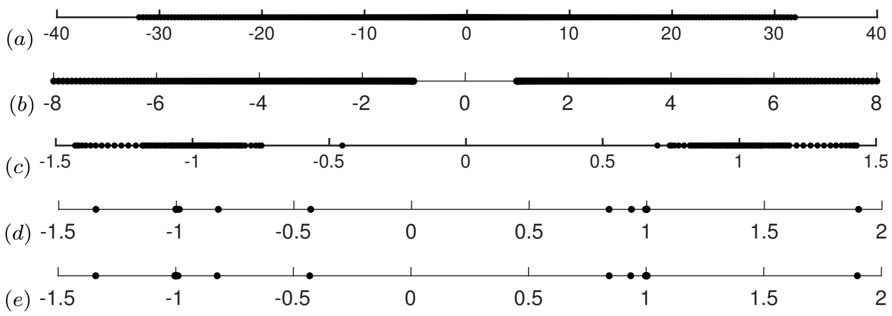

Fig. 6.1 Spectrum of $Y_{512}T_{512}[f]$ with $T_{512}[f]$ generated by $f(x) = x^2 + \mathbf{i}x^3$ (a) with no preconditioner, (b) $T_{512}[z_1]$, (c) $T_{512}[p]$, (d) $P_{512,s}$, or (e) $P_{512,c}$.

Example 6.2 The next example is the ill-conditioned Toeplitz matrix $T_n[f]$ generated by $f(x) = x^4 + \mathbf{i}x^4 \sin x$. In this case,

$$|f(x)| = x^4\sqrt{1 + \sin^2 x}$$

has a zero at $x = 0$ of order 4. Hence, we choose

$$z_2(x) = (2 - 2\cos x)^2$$

of degree $k = 2$.

As this example is severally ill-conditioned, we only test systems with small n. In Tables 6.3 and 6.4, we observe that our proposed band-Toeplitz preconditioners perform well compared with the other preconditioners. In particular, while most preconditioners tested fail, MINRES with $P_{n,c}$ is still effective in the case where $n = 256$. From Figure 6.2, we see the expected spectra around ± 1 with different preconditioners.

Table 6.3 Numbers of MINRES iterations for $Y_nT_n[f]$ with $T_n[f]$ generated by $f(x) = x^4 + \mathbf{i}x^4 \sin x$.

	MINRES						
n	I_n	$\lvert s_n[f]\rvert$	$\lvert c_n[f]\rvert$	$T_n[z_2]$	$T_n[p]$	$P_{n,s}$	$P_{n,c}$
32	95	23	34	35	24	13	16
64	n/a	28	85	73	31	14	16
128	n/a	n/a	n/a	98	37	14	17
256	n/a	n/a	n/a	n/a	n/a	n/a	18

Table 6.4 Numbers of GMRES iterations for $T_n[f]$ generated by $f(x) = x^4 + \mathbf{i}x^4 \sin x$.

	GMRES						
n	I_n	$s_n[f]$	$c_n[f]$	$T_n[z_2]$	$T_n[p]$	$P_{n,s}$	$P_{n,c}$
32	32	24	12	26	15	13	14
64	64	37	14	27	14	12	12
128	128	60	15	24	12	11	11
256	n/a	99	17	20	10	10	10

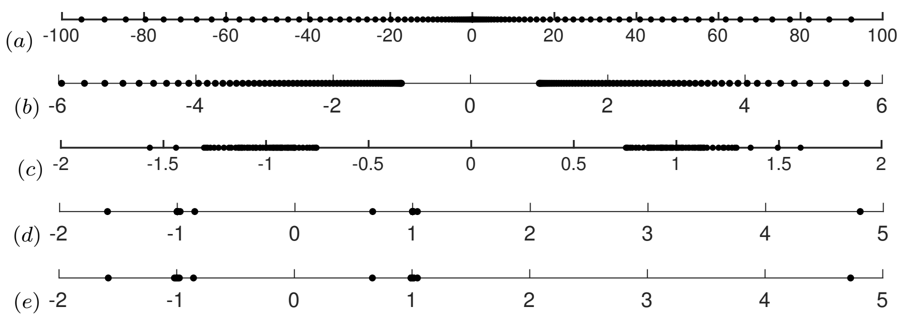

Fig. 6.2 Spectrum of $Y_{128}T_{128}$ with $T_{128}[f]$ generated by $f(x) = x^4 + \mathbf{i}x^4 \sin x$ (a) with no preconditioner, (b) $T_n[z_2]$, (c) $T_n[p]$, (d) $P_{128,s}$, or (e) $P_{128,c}$.

Example 6.3 In this example, we consider the ill-conditioned Toeplitz matrix $T_n[f]$ generated by $f(x) = (x+1)^2(x-1)^2 + \mathbf{i}(x+1)^2(x-1)^2 \sin x$. In this case,

$$|f(x)| = (x+1)^2(x-1)^2\sqrt{1+\sin^2 x}$$

has two distinct zeros at $x = \pm 1$ of order 2. Hence, we pick

$$z_2(x) = \big(2 - 2\cos(x+1)\big)\big(2 - 2\cos(x-1)\big)$$

of degree $k = 2$.

In Tables 6.5 and 6.6, we again observe that our proposed preconditioners, in particular $P_{n,s}$ and $P_{n,c}$, are effective. Surprisingly, while the iteration counts with $|c_n[f]|$ gradually grow with n, $|s_n[f]|$ still performs well in this ill-conditioned example. From Figure 6.3, we again see that the expected spectra with our proposed preconditioners.

Table 6.5 Numbers of MINRES iterations for $Y_nT_n[f]$ with $T_n[f]$ generated by $f(x) = (x+1)^2(x-1)^2 + \mathbf{i}(x+1)^2(x-1)^2 \sin x$.

	MINRES						
n	I_n	$\|s_n[f]\|$	$\|c_n[f]\|$	$T_n[z_2]$	$T_n[p]$	$P_{n,s}$	$P_{n,c}$
512	>1000	17	110	89	39	11	11
1024	>1000	20	155	91	37	11	11
2048	>1000	22	272	91	37	11	11
4096	>1000	20	476	89	36	11	11

Table 6.6 Numbers of GMRES iterations for $T_n[f]$ generated by $f(x) = (x+1)^2(x-1)^2 + \mathbf{i}(x+1)^2(x-1)^2 \sin x$.

	GMRES						
n	I_n	$s_n[f]$	$c_n[f]$	$T_n[z_2]$	$T_n[p]$	$P_{n,s}$	$P_{n,c}$
512	512	8	52	38	22	16	16
1024	>1000	8	70	37	22	16	16
2048	>1000	8	96	36	21	15	15
4096	>1000	7	132	35	20	15	15

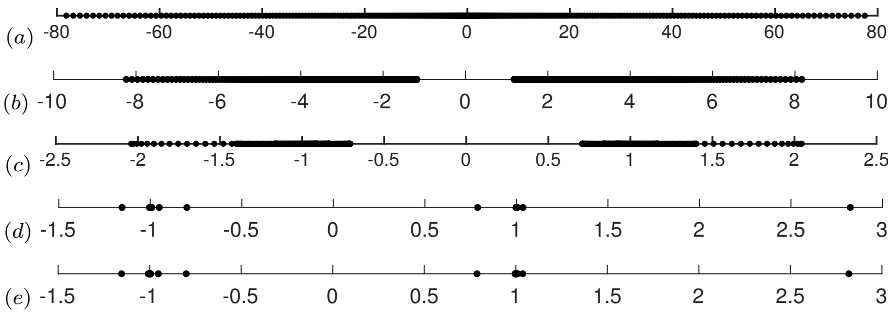

Fig. 6.3 Spectrum of $Y_{512}T_{512}[f]$ with $T_{512}[f]$ generated by $f(x) = (x+1)^2(x-1)^2 + \mathbf{i}(x+1)^2(x-1)^2 \sin x$ (a) with no preconditioner, (b) $T_{512}[z_2]$, (c) $T_{512}[p]$, (d) $P_{512,s}$, or (e) $P_{512,c}$.

Example 6.4 Assuming $m = 2$, we consider the following ill-conditioned block Toeplitz matrix $T_{(n,2)}[f]$ generated by

$$f(x) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x^2 + \mathbf{i}x^3 & 0 \\ 0 & 2 + \cos x \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

In this case, we pick

$$z_1(x) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 - 2\cos x & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

to match the zero of $|f|$.

In Table 6.7, we again observe that our proposed preconditioners, in particular $P_{(n,2;s)}$ and $P_{(n,2;c)}$, are effective. Since absolute circulant preconditioners are not (yet) generalized to the block case, we for this example will consider

MINRES only with our proposed preconditioners. Note that $T_{(n,2)}[p]$ is the pure block band Toeplitz preconditioner based on best approximation, and $P_{(n,2;s)}$ and $P_{(n,2;c)}$ corresponds to the block preconditioners using Strang's circulant preconditioner and optimal circulant preconditioners respectively.

Table 6.8 shows the GMRES iterations as a comparison. The classical block circulant preconditioners (Strang's and T. Chan's) fail to speed up convergence in this ill-conditioned case as expected. Figure 6.4 shows the predicted clustered spectra accordingly to our theorems.

Table 6.7 Numbers of MINRES iterations for $Y_n T_n[f]$.

	MINRES				
$2n$	I_{2n}	$T_{(n,2)}[z_1]$	$T_{(n,2)}[p]$	$P_{(n,2;s)}$	$P_{(n,2;c)}$
64	60	36	22	12	14
128	290	71	27	12	14
256	>1000	103	33	12	14
512	>1000	127	35	12	14

Table 6.8 Numbers of GMRES iterations for $T_n[f]$.

	GMRES						
$2n$	I_{2n}	$s(T_{(n,2)}[f])$	$c(T_{(n,2)}[f])$	$T_{(n,2)}[z_1]$	$T_{(n,2)}[p]$	$P_{(n,2;s)}$	$P_{(n,2;c)}$
64	33	35	37	32	26	22	22
128	77	68	76	48	29	24	24
256	171	134	152	58	28	22	22
512	360	267	300	61	26	21	21

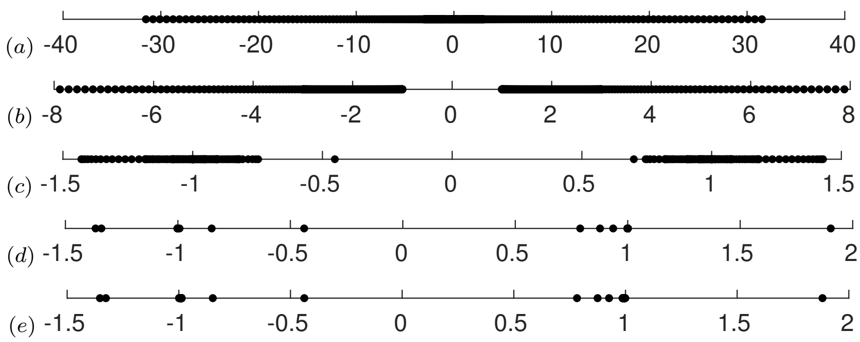

Fig. 6.4 Spectrum of $Y_{512,2}T_{512,2}[f]$ with $T_{512,2}[f]$ as in Example 4 (a) with no preconditioner, (b) $T_{512,2}[z_2]$, (c) $T_{512,2}[p]$, (d) $P_{512,2;s}$, or (e) $P_{512,2;c}$.

Example 6.5 In this example, we apply our proposed method for solving the nonsymmetric Toeplitz system arising from fractional differential equations with constant diffusion coefficients as in (1.2).

Tables 6.9 and 6.10 show the iteration numbers for $Y_nT_n[f]$, where $\alpha = 1.7$, $\nu_{m,n} = \Delta x^{\alpha-1} = (\frac{1}{n+1})^{0.7}$ (i.e. $m = n+1$), $d_+ = 5$, and $d_- = 1$. We construct $T_n[p]$ without considering z_k, and pick $P_{n,s} = s_n[|f|]$ and $P_{n,c} = c_n[|f|]$. The CPU times excluding the time to construct the preconditioners and the exact iteration numbers are also reported. In Figure 6.5, the clustered eigenvalues around ± 1 with our preconditioners are shown when $n = 512$.

Even though the generating function has no even order zeros in this example, the numerical results indicate that both $s_n[|f|]$ and $c_n[|f|]$ are still competitive. Despite its iteration counts are higher, we note that MINRES is faster than GMRES seemingly due to the three-term recurrence in the celebrated Lanczos process. This shows that premultiplying with Y_n is more than just a trick to obtain symmetry for theoretical convergence - it can indeed speed up the solver in practice even in the unpreconditioned case.

Table 6.9 Numbers of MINRES iterations and CPU times (in parenthesis) for $Y_nT_n[f]$ with $T_n[f]$ generated by $f(x) = \nu_{m,n} + d_+g_\alpha(x) + d_-g_\alpha(-x)$.

	MINRES					
n	I_n	$\lvert s_n[f]\rvert$	$\lvert c_n[f]\rvert$	$T_n[p]$	$s_n[\lvert f\rvert]$	$c_n[\lvert f\rvert]$
256	659 (0.068)	12 (0.019)	28 (0.021)	16 (0.007)	11(0.0034)	28 (0.0063)
512	1931 (0.45)	12 (0.015)	30 (0.023)	23 (0.015)	12 (0.0049)	32 (0.016)
1024	5701 (1.18)	12 (0.012)	35 (0.024)	37 (0.031)	12 (0.033)	35 (0.028)
2048	16581 (8.67)	12 (0.027)	35 (0.049)	67 (0.13)	12 (0.013)	36 (0.048)

Table 6.10 Numbers of GMRES iterations and CPU times (in parenthesis) for $Y_nT_n[f]$ with $T_n[f]$ generated by $f(x) = \nu_{m,n} + d_+g_\alpha(x) + d_-g_\alpha(-x)$.

	GMRES					
n	I_n	$s_n[f]$	$c_n[f]$	$T_n[p]$	$s(T_n[\lvert f\rvert])$	$c(T_n[\lvert f\rvert])$
256	239 (0.48)	6 (0.025)	12 (0.020)	11 (0.016)	10 (0.016)	16 (0.016)
512	416 (1.66)	6 (0.019)	12 (0.022)	15 (0.033)	9 (0.018)	16 (0.022)
1024	592 (4.53)	6 (0.033)	12 (0.037)	19 (0.046)	10 (0.041)	15 (0.038)
2048	777 (32.13)	6 (0.054)	11 (0.064)	25 (0.43)	10 (0.061)	14 (0.065)

7 Conclusions

Our proposed preconditioning approach can achieve the optimal $\mathcal{O}(n \log n)$ complexity for a class of nonsymmetric ill-conditioned Toeplitz systems $T_n[f]\mathbf{x}_n =$

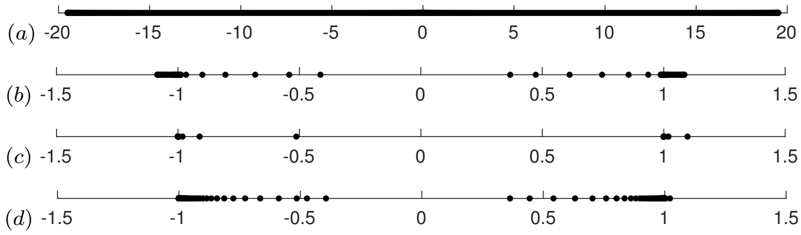

Fig. 6.5 Spectrum of $Y_{512}T_{512}[f]$ with $T_{512}[f]$ generated by $f(x) = \nu_{m,n} + d_+g_\alpha(x) + d_-g_\alpha(-x)$ when $\alpha = 1.7$ (a) with no preconditioner, (b) $T_{512}[p]$, (c) $s(T_{512}[|f|])$ (zoom-in), or (d) $c(T_{512}[|f|])$ (zoom-in).

$\mathbf{b}_n$ without considering normalization. Namely, we have proposed the band-Toeplitz type preconditioners that have theoretical guarantees of effectiveness for $Y_nT_n[f]$ with $T_n[f]$ generated by a continuous, complex-valued function f with even ordered zeros. A number of numerical examples concerning different f were given to support our result. In each ill-conditioned example, our proposed solvers virtually outperformed the others tested, and we have observed significant improvements in convergence and the expected spectra.

Acknowledgements The authors would like to thank the anonymous referees for their constructive comments that improved the original manuscript.

References

1. F. Avram. On bilinear forms in Gaussian random variables and Toeplitz matrices. *Probability Theory and Related Fields*, 79(1):37–45, 1988.
2. D. Bini and F. Di Benedetto. New preconditioner for the parallel solution of positive definite Toeplitz systems. In *Algorithms and Architectures*, SPAA '90, pages 220–223. ACM, 1990.
3. D. Bini and P. Favati. On a Matrix Algebra Related to the Discrete Hartley Transform. *SIAM Journal on Matrix Analysis and Applications*, 14(2):500–507, 1993.
4. A. Böttcher and S. M. Grudsky. *Spectral Properties of Banded Toeplitz Matrices*. Society for Industrial and Applied Mathematics, 2005.
5. S. Serra Capizzano and P. Tilli. Extreme singular values and eigenvalues of non-Hermitian block Toeplitz matrices. *Journal of Computational and Applied Mathematics*, 108(1-2):113–130, 1999.
6. R Chan and K Ng. Toeplitz preconditioners for Hermitian Toeplitz systems. *Linear Algebra and its Applications*, 190:181–208, 1993.
7. R. H. Chan. Toeplitz preconditioners for Toeplitz systems with nonnegative generating functions. *IMA Journal of Numerical Analysis*, 11(3):333–345, 1991.
8. R. H. Chan and M. K. Ng. Conjugate gradient methods for Toeplitz systems. *SIAM Review*, 38(3):427–482, 1996.
9. R. H. Chan, M. K. Ng, and A. M. Yip. The best circulant preconditioners for hermitian Toeplitz systems II: The multiple-zero case. *Numerische Mathematik*, 92(1):17–40, 2002.
10. R. H. Chan and P. T. P. Tang. Fast Band-Toeplitz Preconditioners for Hermitian Toeplitz Systems. *SIAM Journal on Scientific Computing*, 15(1):164–171, 1994.
11. R. H. Chan, A. M. Yip, and N. K. Ng. The best circulant preconditioners for Hermitian Toeplitz systems. *SIAM Journal on Numerical Analysis*, 38(3):876–896, 2001.
12. T. F. Chan. An Optimal Circulant Preconditioner for Toeplitz Systems. *SIAM Journal on Scientific and Statistical Computing*, 9(4):766–771, 1988.

13. F. Di Benedetto, G. Fiorentino, and S. Serra. C. G. preconditioning for Toeplitz matrices. *Computers and Mathematics with Applications*, 25(6):35–45, 1993.
14. M. Donatelli, C. Garoni, M. Mazza, S. Serra-Capizzano, and D. Sesana. Preconditioned HSS method for large multilevel block Toeplitz linear systems via the notion of matrix-valued symbol. *Numerical Linear Algebra with Applications*, 23(1):83–119, 2016.
15. M. Donatelli, M. Mazza, and S. Serra-Capizzano. Spectral analysis and structure preserving preconditioners for fractional diffusion equations. *Journal of Computational Physics*, 307:262–279, 2016.
16. M. Donatelli, M. Neytcheva, and S. Serra-Capizzano. Canonical eigenvalue distribution of multilevel block Toeplitz sequences with non-hermitian symbols. In *Operator Theory: Advances and Applications*, volume 221, pages 269–291, 2012.
17. H. E., D. S., and A. Wathen. *Finite Elements and Fast Iterative Solvers*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, second edition, 2014.
18. Z. W. Fang, M. K. Ng, and H. W. Sun. Circulant preconditioners for a kind of spatial fractional diffusion equations. *Numerical Algorithms*, 82(2):729–747, 2019.
19. P. Ferrari, I. Furci, S. Hon, M. Ayman 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 Journal on Matrix Analysis and Applications*, 40(3):1066–1086, 2019.
20. C. Garoni and S. Serra-Capizzano. *Generalized locally Toeplitz sequences: Theory and applications*, volume 2. Springer, Cham, 2018.
21. S. Hon. Optimal preconditioners for systems defined by functions of Toeplitz matrices. *Linear Algebra and Its Applications*, 548:148–171, 2018.
22. S. Hon. Circulant preconditioners for functions of Hermitian Toeplitz matrices. *Journal of Computational and Applied Mathematics*, 352:328–340, 2019.
23. S. Hon, M. Ayman Mursaleen, and S. Serra-Capizzano. A note on the spectral distribution of symmetrized Toeplitz sequences. *Linear Algebra and Its Applications*, 579(2-3):32–50, 2019.
24. S. Hon and A. Wathen. Circulant preconditioners for analytic functions of Toeplitz matrices. *Numerical Algorithms*, 79(4):1211–1230, 2018.
25. T. Huckle, S. Serra-Capizzano, and C. Tablino-Possio. Preconditioning strategies for non-Hermitian Toeplitz linear systems. *Numerical Linear Algebra with Applications*, 12(2-3):211–220, 2005.
26. X. Q. Jin and Y. M. Wei. A survey and some extensions of T. Chan's preconditioner. *Linear Algebra and Its Applications*, 428(2-3):403–412, 2008.
27. Siu Long Lei and Hai Wei Sun. A circulant preconditioner for fractional diffusion equations. *Journal of Computational Physics*, 2013.
28. M. Mazza and J. Pestana. Spectral properties of flipped Toeplitz matrices and related preconditioning. *BIT Numerical Mathematics*, 59(2):463–482, 2019.
29. M Ng. *Iterative methods for Toeplitz systems*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2004.
30. M Ng. Band preconditioners for block-Toeplitz-Toeplitz-block systems. *Linear Algebra and Its Applications*, 259(1-3):307–327, 1997.
31. D. Noutsos, S. Serra Capizzano, and P. Vassalos. Matrix algebra preconditioners for multilevel Toeplitz systems do not insure optimal convergence rate. *Theoretical Computer Science*, 315(2-3):557–579, 2004.
32. D. Noutsos, G. Tachyridis, Band Toeplitz preconditioners for non-symmetric real Toeplitz systems by preconditioned GMRES method. *Journal of Computational and Applied Mathematics*, 373, 2020.
33. D. Noutsos and P. Vassalos. New band Toeplitz preconditioners for ill-conditioned symmetric positive definite Toeplitz systems. *SIAM Journal on Matrix Analysis and Applications*, 23(3):728–743, 2002.
34. D. Noutsos and P. Vassalos. Superlinear convergence for PCG using band plus algebra preconditioners for Toeplitz systems. *Computers and Mathematics with Applications*, 56(5):1255–1270, 2008.
35. S. V. Parter. On the distribution of the singular values of Toeplitz matrices. *Linear Algebra and Its Applications*, 80(C):115–130, 1986.

36. J. Pestana. Preconditioners for symmetrized Toeplitz and multilevel Toeplitz matrices. *SIAM Journal on Matrix Analysis and Applications*, 40(3):870–887, 2019.
37. J. Pestana and A. J. Wathen. A preconditioned MINRES method for nonsymmetric Toeplitz matrices. *SIAM Journal on Matrix Analysis and Applications*, 36(1):273–288, 2015.
38. L. J. Savage, U. Grenander, and G. Szego. *Toeplitz Forms and Their Applications.*, volume 53. Chelsea Publishing Co., New York, second edition, 1958.
39. S. Serra. Preconditioning strategies for asymptotically ill-conditioned block Toeplitz systems. *Bit*, 34(4):579–594, 1994.
40. S. Serra. Optimal, quasi-optimal and superlinear band-Toeplitz preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems. *Mathematics of Computation*, 66(218):651–666, 1997.
41. S. Serra. Asymptotic results on the spectra of block Toeplitz preconditioned matrices. *SIAM Journal on Matrix Analysis and Applications*, 20(1):31–44, 1998.
42. S. Serra Capizzano and E. Tyrtyshnikov. Any circulant-like preconditioner for multilevel matrices is not superlinear. *SIAM Journal on Matrix Analysis and Applications*, 21(2):431–439, 1999.
43. G. Strang. A proposal for Toeplitz matrix calculations. *Studies in Applied Mathematics*, 74(2):171–176, 1986.
44. E. E. Tyrtyshnikov and N. L. Zamarashkin. Spectra of multilevel Toeplitz matrices: Advanced theory via simple matrix relationships. *Linear Algebra and Its Applications*, 270(1-3):15–27, 1998.
45. E. E. Tyrtyshnikov. A unifying approach to some old and new theorems on distribution and clustering. *Linear Algebra and Its Applications*, 232(1-3):1–43, 1996.
46. A J Wathen. Preconditioning. *Acta Numerica*, 24:329–376, 2015.
47. Z. J. Xie, X. W. Jin, and Y. M. Wei. A fast algorithm for solving circulant tensor systems. *Linear and Multilinear Algebra*, 65(9):1894–1904, 2017.
48. Z. J. Xie, X. Q. Jin, and Z. Zhao. A Convergence Analysis of the MINRES Method for Some Hermitian Indefinite Systems. *East Asian Journal on Applied Mathematics*, 7(4):827–836, 2017.