LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 281 (1998) 265–292

Optimal trigonometric preconditioners for nonsymmetric Toeplitz systems

Daniel Potts [a,1], Gabriele Steidl [b,*]

[a] *Medizinische Universität zu Lübeck, Institut für Mathematik, Wallstr. 40, D-23560 Lübeck, Germany*

[b] *Universität Mannheim, Fakultät für Mathematik und Informatik, D-68131 Mannheim, Germany*

Received 19 January 1997; accepted 16 March 1998

Submitted by G. Heinig

Abstract

This paper is concerned with the solution of systems of linear equations $\boldsymbol{T}_N\boldsymbol{x}_N = \boldsymbol{b}_N$, where $\{\boldsymbol{T}_N\}_{N\in\mathbb{N}}$ denotes a sequence of nonsingular nonsymmetric Toeplitz matrices arising from a generating function of the Wiener class. We present a technique for the fast construction of optimal trigonometric preconditioners $\boldsymbol{M}_N = \boldsymbol{M}_N(\boldsymbol{T}'_N\boldsymbol{T}_N)$ of the corresponding normal equation which can be extended to Toeplitz least squares problems in a straightforward way. Moreover, we prove that the spectrum of the preconditioned matrix $\boldsymbol{M}_N^{-1}\boldsymbol{T}'_N\boldsymbol{T}_N$ is clustered at 1 such that the PCG-method applied to the normal equation converges superlinearly. Numerical tests confirm the theoretical expectations.

AMS classification: 65F10; 65F15; 65T10

Keywords: Toeplitz matrix; Krylov space methods; CG-method; Preconditioners; Normal equation; Clusters of eigenvalues

1. Introduction

Consider the system of linear equations

$$\boldsymbol{T}_N\boldsymbol{x}_N = \boldsymbol{b}_N, \tag{1.1}$$

* Corresponding author. E-mail: steidl@kiwi.math.uni-mannheim.de.

[1] E-mail: potts@informatik.mu-luebeck.de.

PII: S0024-3795(98)10042-3

where $T_N \in \mathbb{R}^{N,N}$ denotes a nonsingular Toeplitz matrix. Toeplitz systems arise in a variety of applications in mathematics and engineering (see [10] and the references therein). While there exist fast and stable direct Toeplitz solvers for Hermitian positive definite Toeplitz matrices T_N, the non-Hermitian case requires additional effort (see e.g. [22]). Iterative methods like GMRES and CG often provide a fast solution of Eq. (1.1) if they are applied in connection with preconditioning techniques [1,10,31]. In particular, these methods profit from the fact that the vector multiplication with the Toeplitz matrix T_N in each iteration step can be computed with $O(N \log N)$ arithmetical operations by using the fast Fourier transform (FFT). Clearly, the multiplication with the preconditioned matrix should have the same or smaller arithmetic complexity. Two types of preconditioners are mainly exploited for linear Toeplitz systems, namely optimal (Cesáro) circulant preconditioners $M_N = C_N(T_N)$ [14] and more simple so-called "Strang" circulant preconditioners $M_N = S_N(T_N)$ [12]. One reason for the choice of circulant preconditioners is the fact that circulant matrices can be diagonalized by the Fourier matrix $F_N := 1/\sqrt{N}\left(e^{-2\pi i jk/N}\right)_{j,k=0}^{N-1}$, where the multiplication of a vector with F_N takes only $O(N \log N)$ arithmetical operations. Moreover, under certain assumptions on the generating function of T_N (see [13,34]), it can be proved that the singular values of $M_N^{-1}T_N$ are clustered at 1. For non-Hermitian T_N, this results in a superlinear convergence of the CG-method applied to the system

$$(M_N^{-1}T_N)^*(M_N^{-1}T_N)x_N = (M_N^{-1}T_N)^*M_N^{-1}b_N. \tag{1.2}$$

To our knowledge, up to now, for non-Hermitian Toeplitz systems and Toeplitz least squares problems only *circulant* preconditioners were constructed. See [13,8] for circulant preconditioners with respect to some kind of normal equation (1.2), [9] for so-called displacement preconditioners and [15]. Note that another approach to the solution of Eq. (1.1) uses relations between Toeplitz-like matrices and Cauchy-like matrices [16,25]. When we finished the paper, we became aware of new results of Tyrtyshnikov et al. [35] concerning the convergence behaviour of the preconditioned GMRES-method which avoids the transition of Eq. (1.1) to the normal equation. However, the preconditioners are again (improved) circulants, which were constructed with respect to T_N.

In this paper, we restrict our attention to *nonsymmetric real* Toeplitz matrices T_N. Here, it seems to be natural, to replace the circulant matrices by matrices which are diagonalizable by some real trigonometric matrices. In this way, arithmetic with complex numbers can be completely avoided. Of course, the commonly used trigonometric transforms are closely related to the Fourier transform. Indeed, for *symmetric* Toeplitz matrices T_N with positive continuous 2π-periodic generating functions, trigonometric preconditioning significantly accelates the convergence of the CG-method (see [3–5, 7,11,19,24]).

In this paper, we suggest the solution of Eq. (1.1) by applying the CG-method to the preconditioned normal equation

$$M_N^{-1} T_N' T_N x_N = M_N^{-1} T_N' b_N,$$

where in contrast to Eq. (1.2), $M_N = M_N(T_N' T_N)$ denotes the optimal preconditioner with respect to $T_N' T_N$. We demonstrate that the construction of such optimal preconditioners can be realized with only $O(N \log N)$ arithmetical operations despite the fact that $T_N' T_N$ is no longer a Toeplitz matrix. We prove that under certain assumptions on T_N the eigenvalues of $M_N^{-1} T_N' T_N$ are clustered at 1. Although our approach works in exactly the same way for different trigonometric transforms, we prefer to investigate the DCT-II preconditioner in detail and add only few facts concerning the other trigonometric preconditioners. We hope that our notation makes the approach for other trigonometric preconditioners immediately clear. Numerical tests were performed for the different trigonometric preconditioners. In all examples our preconditioning was superior over the method (1.2) with an optimal trigonometric preconditioner $M_N(T_N)$ of T_N.

This paper is organized as follows: Section 2 contains the basic matrix notation. In Section 3, we study the relations between trigonometric transforms and Toeplitz matrices. In particular, we introduce a method for the fast vector multiplication with real nonsymmetric Toeplitz matrices based on real trigonometric transforms. In Section 4, we introduce optimal trigonometric preconditioners. Section 5 is concerned with the proof that the eigenvalues of the preconditioned matrix $M_N^{-1} T_N' T_N$ are clustered at 1. In Section 6, we present the fast construction of the optimal preconditioner. Finally, Section 7 confirms the theoretical expectations by numerical tests.

2. Notation

For the sake of clarity, we collect the matrix notation in this preliminary section. Let $a_N := (a_0, \ldots, a_{N-1})'$, $b_N := (b_0, \ldots, b_{N-1})'$ and let o_N be the vector consisting of N zeros. Here A' is the transpose of A. By I_N we denote the (N,N)-identity matrix and by $e_k \in \mathbb{R}^N$ the kth identity vector. To describe Toeplitz and Hankel matrices, we use the following notation:

$$\operatorname{toeplitz}(a', b') := \begin{pmatrix} a_0 & a_1 & \ldots & a_{N-2} & a_{N-1} \\ b_1 & a_0 & \ldots & a_{N-3} & a_{N-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{N-2} & b_{N-3} & \ldots & a_0 & a_1 \\ b_{N-1} & b_{N-2} & \ldots & b_1 & a_0 \end{pmatrix} \quad (\text{with } a_0 = b_0),$$

stoeplitz a': symmetric Toeplitz matrix with first row a', atoeplitz $\boldsymbol{a}'$: antisymmetric Toeplitz matrix with first row $\boldsymbol{a}'$, where $a_0 = 0$,

$$\text{hankel}(\boldsymbol{a}', \boldsymbol{b}') := \begin{pmatrix} a_0 & a_1 & \dots & a_{N-2} & a_{N-1} \\ a_1 & a_2 & \dots & a_{N-1} & b_{N-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{N-2} & a_{N-1} & \dots & b_2 & b_1 \\ a_{N-1} & b_{N-2} & \dots & b_1 & b_0 \end{pmatrix} \quad (\text{with } a_{N-1} = b_{N-1}),$$

shankel $\boldsymbol{a}'$: persymmetric Hankel matrix with first row $\boldsymbol{a}'$,
ahankel $\boldsymbol{a}'$: antipersymmetric Hankel matrix with first row $\boldsymbol{a}'$, where $a_{N-1} = 0$.
Further, we introduce the matrices

$$\boldsymbol{Z}'_{N,1} := \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{N,N+1},$$

$$\boldsymbol{Z}'_{N,2} = \begin{pmatrix} 1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & 0 \end{pmatrix} \in \mathbb{R}^{N,N+1}$$

and

$$\boldsymbol{R}'_N := \begin{pmatrix} 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \in \mathbb{R}^{N-1,N+1}.$$

Let diag $\boldsymbol{a}$ be the diagonal matrix with diagonal $\boldsymbol{a}$ and let $\delta(\boldsymbol{A}) := \text{diag}(a_{k,k})_{k=0}^{N-1}$, where $a_{k,k}$ is the (k,k)th entry of $\boldsymbol{A}$. By

$$\text{tr}\, \boldsymbol{A} := \sum_{k=0}^{N-1} a_{k,k},$$

we denote the *trace* of $\boldsymbol{A}$. Moreover, we need the following *matrix norms*:
Spectral norm:

$$||\boldsymbol{A}||_2 := \text{ (maximum of the singular values of } \boldsymbol{A})^{1/2},$$

Frobenius norm:

$$||\boldsymbol{A}||_F := \left(\sum_{j,k=0}^{N-1} a_{jk}^2 \right)^{1/2},$$

1-*norm*:

$$\|A\|_1 := \max\left\{\sum_{j=0}^{N-1} |a_{j,k}| : k = 0, \ldots, N-1\right\}.$$

If it does not make confusion, we use the same notation for the norm of absolute summable sequences $a = \{a_k\}_{k\in\mathbb{Z}} \in l_1$, i.e.

$$\|a\|_1 := \sum_{k\in\mathbb{Z}} |a_k|.$$

3. Trigonometric transforms and Toeplitz matrices

We introduce four discrete sine transforms (DST) and four discrete cosine transforms (DCT) as classified by Wang [36]:

$$\begin{aligned}
\text{DCT-I}: \ \boldsymbol{C}^{I}_{N+1} &:= \left(\frac{2}{N}\right)^{1/2} \left(\varepsilon^N_j \varepsilon^N_k \cos\frac{jk\pi}{N}\right)_{j,k=0}^{N} \in \mathbb{R}^{N+1,N+1},\\
\text{DCT-II}: \ \boldsymbol{C}^{II}_{N} &:= \left(\frac{2}{N}\right)^{1/2} \left(\varepsilon^N_j \cos\frac{j(2k+1)\pi}{2N}\right)_{j,k=0}^{N-1} \in \mathbb{R}^{N,N},\\
\text{DCT-III}: \ \boldsymbol{C}^{III}_{N} &:= (\boldsymbol{C}^{II}_N)' \in \mathbb{R}^{N,N},\\
\text{DCT-IV}: \ \boldsymbol{C}^{IV}_{N} &:= \left(\frac{2}{N}\right)^{1/2} \left(\cos\frac{(2j+1)(2k+1)\pi}{4N}\right)_{j,k=0}^{N-1} \in \mathbb{R}^{N,N}
\end{aligned}$$

and

$$\begin{aligned}
\text{DST-I}: \ \boldsymbol{S}^{I}_{N-1} &:= \left(\frac{2}{N}\right)^{1/2} \left(\sin\frac{(j+1)(k+1)\pi}{N}\right)_{j,k=0}^{N-2} \in \mathbb{R}^{N-1,N-1},\\
\text{DST-II}: \ \boldsymbol{S}^{II}_{N} &:= \left(\frac{2}{N}\right)^{1/2} \left(\varepsilon^N_{j+1} \sin\frac{(j+1)(2k+1)\pi}{2N}\right)_{j,k=0}^{N-1} \in \mathbb{R}^{N,N},\\
\text{DST-III}: \ \boldsymbol{S}^{III}_{N} &:= (\boldsymbol{S}^{II}_N)' \in \mathbb{R}^{N,N},\\
\text{DST-IV}: \ \boldsymbol{S}^{IV}_{N} &:= \left(\frac{2}{N}\right)^{1/2} \left(\sin\frac{(2j+1)(2k+1)\pi}{4N}\right)_{j,k=0}^{N-1} \in \mathbb{R}^{N,N},
\end{aligned}$$

where $\varepsilon^N_k := 1/\sqrt{2}$ $(k = 0, N)$ and $\varepsilon^N_k := 1$, otherwise. We refer to the corresponding transforms as *trigonometric transforms*. It is well-known that the above matrices are orthogonal and that the vector multiplication with any of these matrices takes only $O(N \log N)$ arithmetical operations. Fortunately, there exist implementations of algorithms for the vector multiplication with the above sine and cosine matrices, for example a C-implementation based on [2,31].

Moreover, we use the slightly modified DCT-I and DST-I matrices

$$\tilde{C}^I_{N+1} := \left((\varepsilon_k^N)^2 \cos\frac{jk\pi}{N}\right)_{j,k=0}^{N}, \qquad \tilde{S}^I_{N-1} := \left(\sin\frac{jk\pi}{N}\right)_{j,k=1}^{N-1}$$

and the slightly modified DCT-III and DST-III matrices

$$\tilde{C}^{III}_N := \left((\varepsilon_k^N)^2 \cos\frac{(2j+1)k\pi}{2N}\right)_{j,k=0}^{N-1},$$

$$\tilde{S}^{III}_N := \left((\varepsilon_{k+1}^N)^2 \sin\frac{(2j+1)(k+1)\pi}{2N}\right)_{j,k=0}^{N-1}.$$

Then

$$\tilde{C}^I_{N+1}\tilde{C}^I_{N+1} = \frac{N}{2} I_{N+1}. \tag{3.1}$$

Theorem 3.1. *There exist the following relations between trigonometric transforms and Toeplitz matrices:*

(i) *DCT-I and DST-I*:

$$\begin{aligned}
R_N' C^I_{N+1}\, D\, C^I_{N+1} R_N &= \tfrac{1}{2}\,\mathrm{stoeplitz}(a_0,\ldots,a_{N-2}) \\
&\quad + \tfrac{1}{2}\,\mathrm{shankel}(a_2,\ldots,a_{N-2},0,0), \\
S^I_{N-1}\, R_N'\, D\, R_N S^I_{N-1} &= \tfrac{1}{2}\,\mathrm{stoeplitz}(a_0,\ldots,a_{N-2}) \\
&\quad - \tfrac{1}{2}\,\mathrm{shankel}(a_2,\ldots,a_{N-2},0,0), \\
R_N' C^I_{N+1}\, \tilde{D}\, R_N S^I_{N-1} &= \tfrac{1}{2}\,\mathrm{atoeplitz}(0,a_1,\ldots,a_{N-2}) \\
&\quad + \tfrac{1}{2}\,\mathrm{ahankel}(a_2,\ldots,a_{N-1},0), \\
S^I_{N-1}\, R_N'\, \tilde{D}\, C^I_{N+1} R_N &= -\tfrac{1}{2}\,\mathrm{atoeplitz}(0,a_1,\ldots,a_{N-2}) \\
&\quad + \tfrac{1}{2}\,\mathrm{ahankel}(a_2,\ldots,a_{N-1},0)
\end{aligned}$$

with

$$\begin{aligned}
&D := \mathrm{diag}(d_0,\ldots,d_N)', \qquad \tilde{D} := \mathrm{diag}(0,\tilde{d}_1,\ldots,\tilde{d}_{N-1},0)', \\
&(d_0,\ldots,d_N)' := \tilde{C}^I_{N+1}\,(a_0,\ldots,a_{N-2},0,0)', \\
&(\tilde{d}_1,\ldots,\tilde{d}_{N-1})' := \tilde{S}^I_{N-1}\,(a_1,\ldots,a_{N-1})'.
\end{aligned}$$

(ii) *DCT-II and DST-II*:

$$
\begin{aligned}
(\boldsymbol{C}_N^{II})'\, \boldsymbol{Z}_{N,2}'\, \boldsymbol{D}\, \boldsymbol{Z}_{N,2} \boldsymbol{C}_N^{II} &= \tfrac{1}{2}\,\mathrm{stoeplitz}(a_0,\ldots,a_{N-1}) \\
&\quad + \tfrac{1}{2}\,\mathrm{shankel}(a_1,\ldots,a_{N-1},0), \\
(\boldsymbol{S}_N^{II})'\, \boldsymbol{Z}_{N,1}'\, \boldsymbol{D}\, \boldsymbol{Z}_{N,1} \boldsymbol{S}_N^{II} &= \tfrac{1}{2}\,\mathrm{stoeplitz}(a_0,\ldots,a_{N-1}) \\
&\quad - \tfrac{1}{2}\,\mathrm{shankel}(a_1,\ldots,a_{N-1},0), \\
(\boldsymbol{C}_N^{II})'\, \boldsymbol{Z}_{N,2}'\, \tilde{\boldsymbol{D}}\, \boldsymbol{Z}_{N,1} \boldsymbol{S}_N^{II} &= \tfrac{1}{2}\,\mathrm{atoeplitz}(0,a_1,\ldots,a_{N-1}) \\
&\quad + \tfrac{1}{2}\,\mathrm{ahankel}(a_1,\ldots,a_{N-1},0), \\
(\boldsymbol{S}_N^{II})'\, \boldsymbol{Z}_{N,1}'\, \tilde{\boldsymbol{D}}\, \boldsymbol{Z}_{N,2} \boldsymbol{C}_N^{II} &= -\tfrac{1}{2}\,\mathrm{atoeplitz}(0\,a_1,\ldots,a_{N-1}) \\
&\quad + \tfrac{1}{2}\,\mathrm{ahankel}(a_1,\ldots,a_{N-1},0)
\end{aligned}
$$

with

$$
\begin{aligned}
&\boldsymbol{D} := \mathrm{diag}(d_0,\ldots,d_N)', \qquad \tilde{\boldsymbol{D}} := \mathrm{diag}(0,\tilde{d}_1,\ldots,\tilde{d}_{N-1},0)', \\
&(d_0,\ldots,d_N)' := \tilde{\boldsymbol{C}}_{N+1}^{I}\,(a_0,\ldots,a_{N-1},0)', \\
&(\tilde{d}_1,\ldots,\tilde{d}_{N-1})' := \tilde{\boldsymbol{S}}_{N-1}^{I}\,(a_1,\ldots,a_{N-1})'.
\end{aligned}
$$

(iii) *DCT-IV and DST-IV*:

$$
\begin{aligned}
\boldsymbol{C}_N^{IV}\, \boldsymbol{D}\, \boldsymbol{C}_N^{IV} &= \tfrac{1}{2}\,\mathrm{stoeplitz}(a_0,\ldots,a_{N-1}) + \tfrac{1}{2}\,\mathrm{ahankel}(a_1,\ldots,a_{N-1},0), \\
\boldsymbol{S}_N^{IV}\, \boldsymbol{D}\, \boldsymbol{S}_N^{IV} &= \tfrac{1}{2}\,\mathrm{stoeplitz}(a_0,\ldots,a_{N-1}) - \tfrac{1}{2}\,\mathrm{ahankel}(a_1,\ldots,a_{N-1},0), \\
\boldsymbol{C}_N^{IV}\, \tilde{\boldsymbol{D}}\, \boldsymbol{S}_N^{IV} &= \tfrac{1}{2}\,\mathrm{atoeplitz}(0,a_1,\ldots,a_{N-1}) + \tfrac{1}{2}\,\mathrm{shankel}(a_1,\ldots,a_{N-1},0), \\
\boldsymbol{S}_N^{IV}\, \boldsymbol{D}\, \boldsymbol{C}_N^{IV} &= -\tfrac{1}{2}\,\mathrm{atoeplitz}(0,a_1,\ldots,a_{N-1}) + \tfrac{1}{2}\,\mathrm{shankel}(a_1,\ldots,a_{N-1},0)
\end{aligned}
$$

with

$$
\begin{aligned}
&\boldsymbol{D} := \mathrm{diag}(d_0,\ldots,d_{N-1})', \qquad \tilde{\boldsymbol{D}} := \mathrm{diag}(\tilde{d}_0,\ldots,\tilde{d}_{N-1})', \\
&(d_0,\ldots,d_{N-1})' := \boldsymbol{C}_N^{III}\,(a_0,\ldots,a_{N-1})', \\
&(\tilde{d}_0,\ldots,\tilde{d}_{N-1})' := \tilde{\boldsymbol{S}}_N^{III}\,(a_1,\ldots,a_{N-1},0)'.
\end{aligned}
$$

Note that for fixed diagonal matrices $\boldsymbol{D}, \tilde{\boldsymbol{D}}$, the above decompositions into a Toeplitz and a Hankel matrix are not unique.

Proof. We restrict the proof to the DCT-II. To simplify the notation, we drop the index N and set $\boldsymbol{C} := \boldsymbol{Z}_{N,2}\boldsymbol{C}_N^{II}$ and $\boldsymbol{D} := \mathrm{diag}(d_0,\ldots,d_N)'$. Then by

$$\cos\alpha\,\cos\beta = \tfrac{1}{2}\cos(\alpha-\beta) + \tfrac{1}{2}\cos(\alpha+\beta),$$

the (u,v)-entry of the matrix $\boldsymbol{C}'\boldsymbol{D}\boldsymbol{C}$ is

$$(\boldsymbol{C}'\boldsymbol{D}\boldsymbol{C})_{u,v} = \frac{1}{2}\,\frac{2}{N}\sum_{k=0}^{N-1} (\varepsilon_k^N)^2\, d_k \cos\frac{(u-v)k\pi}{N} + \frac{1}{2}\,\frac{2}{N}\sum_{k=0}^{N-1} (\varepsilon_k^N)^2\, d_k \cos\frac{(u+v+1)k\pi}{N},$$

or equivalently, since $-(-1)^{u-v}d_N = (-1)^{u+v+1}d_N$ for arbitrary $d_N \in \mathbb{R}$,

$$(\boldsymbol{C}'\boldsymbol{D}\boldsymbol{C})_{u,v} = \frac{1}{2}\,\frac{2}{N}\sum_{k=0}^{N} (\varepsilon_k^N)^2\, d_k \cos\frac{(u-v)k\pi}{N} + \frac{1}{2}\,\frac{2}{N}\sum_{k=0}^{N} (\varepsilon_k^N)^2\, d_k \cos\frac{(u+v+1)k\pi}{N}.$$

Choosing $d_N \in \mathbb{R}$ such that

$$\sum_{k=0}^{N} (\varepsilon_k^N)^2\, d_k(-1)^k = 0,$$

we get by symmetry properties of cosine function that

$$\boldsymbol{C}'\boldsymbol{D}\,\boldsymbol{C} = \tfrac{1}{2}\,\mathrm{stoeplitz}(a_0,\ldots,a_{N-1}) + \tfrac{1}{2}\,\mathrm{shankel}(a_1,\ldots,a_{N-1},0),$$

where

$$(a_0,\ldots,a_{N-1},0)' = \frac{2}{N}\,\tilde{\boldsymbol{C}}^I_{N+1}\,(d_0,\ldots,d_N)',$$

i.e. by Eq. (3.1)

$$(d_0,\ldots,d_N)' = \tilde{\boldsymbol{C}}^I_{N+1}\,(a_0,\ldots,a_{N-1},0)'.$$

The other decomposition relations follow in a similar way by application of $\sin\alpha\sin\beta = \frac{1}{2}\cos(\alpha-\beta) - \frac{1}{2}\cos(\alpha+\beta)$ and $\sin\alpha\cos\beta = \frac{1}{2}\sin(\alpha-\beta) + \frac{1}{2}\sin(\alpha+\beta)$. □

Theorem 3.1 provides a new method for the fast multiplication of a real vector with a real nonsymmetric Toeplitz matrix that avoids the complex arithmetic which comes into the play if we exploit the usual FFT-based method for the fast vector – Toeplitz matrix multiplication.

Corollary 3.2 (Fast vector multiplication with nonsymmetric Toeplitz matrices). *Let*

$$\boldsymbol{T} = \boldsymbol{T}_N := \left(t_{j-k}\right)_{j,k=0}^{N-1} = \mathrm{toeplitz}((t_0,t_{-1},\ldots,t_{-(N-1)}),(t_0,t_1,\ldots,t_{N-1}))$$

be given and let $\boldsymbol{C} := \boldsymbol{Z}_{N,2}\boldsymbol{C}_N^{II}$, $\boldsymbol{S} := \boldsymbol{Z}_{N,1}\boldsymbol{S}_N^{II}$. *Then*

$$\boldsymbol{T} = \tfrac{1}{2}(\boldsymbol{T}+\boldsymbol{T}') + \tfrac{1}{2}(\boldsymbol{T}-\boldsymbol{T}') = \boldsymbol{C}'\boldsymbol{D}\boldsymbol{C} + \boldsymbol{S}'\boldsymbol{D}\boldsymbol{S} + \boldsymbol{C}'\tilde{\boldsymbol{D}}\boldsymbol{S} - \boldsymbol{S}'\tilde{\boldsymbol{D}}\boldsymbol{C},$$

where

$$\boldsymbol{D} := \operatorname{diag}(d_0,\ldots,d_N)', \qquad \tilde{\boldsymbol{D}} := \operatorname{diag}(0,\tilde{d}_1,\ldots,\tilde{d}_{N-1},0)',$$

$$(d_0,\ldots,d_N)' := \tilde{\boldsymbol{C}}_{N+1}^{I}\left(t_0,\frac{t_1+t_{-1}}{2},\ldots,\frac{t_{N-1}+t_{-(N-1)}}{2},0\right)',$$

$$(\tilde{d}_1,\ldots,\tilde{d}_{N-1})' := \tilde{\boldsymbol{S}}_{N-1}^{I}\left(\frac{t_{-1}-t_1}{2},\ldots,\frac{t_{-(N-1)}-t_{N-1}}{2}\right)'.$$

The vector multiplication with $\boldsymbol{T}$ *requires except of* $\mathrm{O}(N)$ *additions*
one DCT-I and one DST-I to build $\boldsymbol{D}$ *and* $\tilde{\boldsymbol{D}}$ *in a precomputation step,*
one DCT-II and one DST-II,
four multiplications of vectors with diagonal matrices,
one DCT-III and one DST-III of the vectors $\boldsymbol{DCx} + \tilde{\boldsymbol{D}}\boldsymbol{Sx}$ *and* $\boldsymbol{DSx} - \tilde{\boldsymbol{D}}\boldsymbol{Cx}$,
respectively, and takes therefore only $\mathrm{O}(N \log N)$ *arithmetical operations.*

Clearly, by Theorem 3.1, we can formulate similar algorithms for the fast multiplication of vectors with Toeplitz or Hankel matrices with respect to the other trigonometric transforms. Typewriting this paper, we got a ps-file of a paper of Heinig and Rost [18], which contains results in a similar direction as presented in this section. Following the hint of one of the referees, we realized that representations of symmetric Toeplitz matrices based on DCT-I and DST-I were also given in Refs. [20,21,23].

4. Optimal trigonometric preconditioners

We are concerned with the solution of the system of linear equations

$$\boldsymbol{T}_N\boldsymbol{x}_N = \boldsymbol{b}_N$$

with a nonsingular nonsymmetric Toeplitz matrix $\boldsymbol{T}_N \in \mathbb{R}^{N,N}$. We intend to solve the normal equation

$$\boldsymbol{T}_N'\boldsymbol{T}_N\boldsymbol{x}_N = \boldsymbol{T}_N'\boldsymbol{b}_N \tag{4.1}$$

by the CG-method. In Section 7, we will see that with a good preconditioner at hand, this can be realized in a fast way. There are several requirements on a preconditioner $\boldsymbol{M}_N$ of Eq. (4.1) resulting from the construction and the convergence behaviour of the CG-method as well as from the fact that the vector multiplication with $\boldsymbol{T}_N$ requires only $\mathrm{O}(N \log N)$ arithmetical operations. Therefore, we are looking for a preconditioner with the following properties:

(P1) $\boldsymbol{M}_N$ is symmetric and positive definite such that the bilinear form

$$(\boldsymbol{x}_N, \boldsymbol{y}_N)_{\boldsymbol{M}_N} := \boldsymbol{x}'_N \boldsymbol{M}_N \boldsymbol{y}_N$$

arising in the left preconditioned CG-method is symmetric and positive definite, too.

(P2) The spectrum of $\boldsymbol{M}_N^{-1} \boldsymbol{T}'_N \boldsymbol{T}_N$ is clustered at 1.

(P3) The vector multiplication with $\boldsymbol{M}_N$ can be computed with $\mathrm{O}(N \log N)$ arithmetical operations.

(P4) The construction of $\boldsymbol{M}_N$ takes only $\mathrm{O}(N \log N)$ arithmetical operations.

Having property (P3) in mind, a straightforward idea consists in choosing $\boldsymbol{M}_N$ from an algebra

$$\mathscr{A}_{\boldsymbol{O}_N} := \{\boldsymbol{O}'_N (\operatorname{diag} \boldsymbol{d}) \boldsymbol{O}_N : \boldsymbol{d} \in \mathbb{R}^N\} \tag{4.2}$$

of matrices which are diagonalizable by some orthogonal matrix $\boldsymbol{O}_N$, where $\boldsymbol{O}_N$ has the additional property that its vector multiplication requires only $\mathrm{O}(N \log N)$ arithmetical operations. As orthogonal matrices, we will use the trigonometric matrices of the previous section which are closely related to the Fourier matrix $\boldsymbol{F}_N$, but have the advantage of purely real entries. Moreover, if we choose $\boldsymbol{M}_N \in \mathscr{A}_{\boldsymbol{O}_N}$ as so-called optimal preconditioner of $\boldsymbol{T}'_N \boldsymbol{T}_N$, then we will see that under certain assumptions on $\boldsymbol{T}_N$, the properties (P1), (P2) and (P4) are also fulfilled. For $\boldsymbol{A}_N \in \mathbb{R}^{N,N}$, the matrix $\boldsymbol{M}_N(\boldsymbol{A}_N)$ is called an *optimal preconditioner* of $\boldsymbol{A}_N$ in $\mathscr{A}_{\boldsymbol{O}_N}$ [14] if

$$\|\boldsymbol{M}_N(\boldsymbol{A}_N) - \boldsymbol{A}_N\|_F = \min\{\|\boldsymbol{B}_N - \boldsymbol{A}_N\|_F : \boldsymbol{B}_N \in \mathscr{A}_{\boldsymbol{O}_N}\}. \tag{4.3}$$

If $\boldsymbol{O}_N$ is one of the orthogonal matrices which correspond to the DST-I, DST-II, DCT-II, DST-IV or DCT-IV, respectively, then $\boldsymbol{M}_N$ is said to be an *optimal trigonometric preconditioner* of $\boldsymbol{A}_N$. The choice of the Frobenius norm in definition (4.3) results from the fact that the Frobenius norm is induced by an inner product of $\mathbb{R}^{N,N}$

$$\langle \boldsymbol{A}_N, \boldsymbol{B}_N \rangle := \operatorname{tr}(\boldsymbol{A}'_N \boldsymbol{B}_N) = \sum_{j,k=0}^{N-1} a_{j,k} b_{j,k}. \tag{4.4}$$

In particular, we have

$$\|\boldsymbol{O}_N \boldsymbol{A}_N \boldsymbol{O}'_N\|_F^2 = \operatorname{tr}(\boldsymbol{O}_N \boldsymbol{A}'_N \boldsymbol{O}'_N \boldsymbol{O}_N \boldsymbol{A}_N \boldsymbol{O}'_N) = \operatorname{tr}(\boldsymbol{A}'_N \boldsymbol{A}_N) = \|\boldsymbol{A}_N\|_F^2. \tag{4.5}$$

The following lemma describes the optimal preconditioner of $\boldsymbol{T}'_N \boldsymbol{T}_N$ in two different ways.

Lemma 4.1. *Let $\boldsymbol{A}_N \in \mathbb{R}^{N,N}$ and let $\mathscr{A}_{\boldsymbol{O}_N}$ be defined by Eq. (4.2) with respect to some orthogonal matrix $\boldsymbol{O}_N$. Then the optimal preconditioner of $\boldsymbol{A}_N$ is given by*

$$\boldsymbol{M}_N(\boldsymbol{A}_N) = \boldsymbol{O}'_N \delta(\boldsymbol{O}_N \boldsymbol{A}_N \boldsymbol{O}'_N) \boldsymbol{O}_N. \tag{4.6}$$

If $\{\boldsymbol{B}_0, \ldots, \boldsymbol{B}_{N-1}\}$ *denotes a basis of* $\mathscr{A}_{O_N}$, *then an alternative description of* $\boldsymbol{M}_N(\boldsymbol{A}_N)$ *reads as*

$$\boldsymbol{M}_N(\boldsymbol{A}_N) := \sum_{k=0}^{N-1} \alpha_k \boldsymbol{B}_k, \tag{4.7}$$

where the coefficient vector $\boldsymbol{\alpha} := (\alpha_0, \ldots, \alpha_{N-1})'$ *is determined by*

$$\boldsymbol{G}\boldsymbol{\alpha} = \boldsymbol{\beta}, \qquad \boldsymbol{G} := \left(\langle \boldsymbol{B}_k, \boldsymbol{B}_j \rangle\right)_{j,k=0}^{N-1}, \qquad \boldsymbol{\beta} := \left(\langle \boldsymbol{A}_N, \boldsymbol{B}_j \rangle\right)_{j=0}^{N-1}.$$

Proof. 1. By Eq. (4.5), it follows for $\boldsymbol{M}_N := \boldsymbol{O}'_N(\operatorname{diag} \boldsymbol{d})\boldsymbol{O}_N \in \mathscr{A}_{O_N}$ that

$$||\boldsymbol{M}_N - \boldsymbol{A}_N||_F = ||\operatorname{diag} \boldsymbol{d} - \boldsymbol{O}_N \boldsymbol{A}_N \boldsymbol{O}'_N||_F$$

which implies Eq. (4.6) by the definition of the optimal preconditioner.

2. The computation of the optimal preconditioner of $\boldsymbol{A}_N$ in $\mathscr{A}_{O_N}$ is equivalent with the computation of the element of best approximation of $\boldsymbol{A}_N$ in the linear subspace $\mathscr{A}_{O_N}$ of the Hilbert space $\mathbb{R}^{N,N}$ equipped with the inner product (4.4). This can be done by the Galerkin approach (4.7). □

Now it is easy to verify that an optimal preconditioner of $\boldsymbol{T}'_N \boldsymbol{T}_N$ satisfies property (P1).

Corollary 4.2. *Let* $A_N \in R^{N,N}$ *be a symmetric positive definite matrix and let* $\mathscr{A}_{O_N}$ *be defined by Eq.* (4.2) *with respect to some orthogonal matrix* $\boldsymbol{O}_N$. *Then the optimal preconditioner* $\boldsymbol{M}_N = \boldsymbol{M}_N(\boldsymbol{A}_N)$ *is also symmetric and positive definite.*

For the proof of Corollary 4.2 and in connection with Eq. (4.6) see for example [32,20].

5. Clusters of eigenvalues

Let $C_{2\pi}$ denote the Banach space of 2π-periodic complex-valued functions equipped with the usual norm $||\cdot||_\infty$. In this section, we are only interested in functions $f = f_R + \mathrm{i} f_I \in C_{2\pi}$ with real Fourier coefficients

$$t_k := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \mathrm{e}^{-\mathrm{i}kx} \, \mathrm{d}x \quad (k \in \mathbb{Z}),$$

where we suppose that the real and the imaginary part of f

$$f_{\mathrm{R}} = t_0 + \sum_{k=1}^{\infty}(t_k + t_{-k}) \cos kx,$$

$$f_{\mathrm{I}} = \sum_{k=1}^{\infty}(t_k - t_{-k}) \sin kx$$

do not vanish, respectively. Moreover, we assume that $\{t_k\}_{k\in\mathbb{Z}} \in l_1$, such that f belongs to the Wiener class. Consider the *Nth Toeplitz matrix corresponding to the generating function f*

$$\boldsymbol{T}_N := \mathrm{toeplitz}((t_0, t_{-1}, \ldots, t_{-(N-1)}), (t_0, t_1, \ldots, t_{N-1})).$$

It is well-known that the singular values of T_N are distributed as $|f|$ [27]. Note that the above result was extended to functions $f \in L^2_{2\pi} \supset C_{2\pi}$ in [34].

The following definition is due to Tyrtyshnikov [34]. Let $\{\sigma_k^N\}_{k=1}^N$ be a sequence of real numbers and let $\gamma_N(\varepsilon)$ denote the number of those among σ_k^N $(k = 1, \ldots, N)$ which are outside the ε-ball centred at p. If $\gamma_N(\varepsilon) < K(\varepsilon)$, where $K(\varepsilon)$ is independent of N, then p is called a *proper cluster*. In this sense, we say that the values σ_k^N are clustered at p.

In the following, we restrict our attention to preconditioners of $\mathscr{A}_{C_N^{II}}$. By Theorem 3.1, the approach for the preconditioners with respect to the DST-I, the DST-II, the DST-IV and the DCT-IV follows the same lines. Let $\boldsymbol{C} = \boldsymbol{C}_N := \boldsymbol{Z}_{N,2}\boldsymbol{C}_N^{II}$ and $\boldsymbol{S} = \boldsymbol{S}_N := \boldsymbol{Z}_{N,1}\boldsymbol{S}_N^{II}$. Regarding Theorem 3.1(ii), we associate with the sequence $\{\boldsymbol{T}_N\}_{N\in\mathbb{N}}$ of Toeplitz matrices a sequence $\{\boldsymbol{H}_N\}_{N\in\mathbb{N}}$ of Hankel matrices

$$\boldsymbol{H}_N := \mathrm{hankel}((t_{-1}, \ldots, t_{-(N-1)}, 0), (t_1, \ldots, t_{N-1}, 0)).$$

Let $\boldsymbol{M}_N$ denote the optimal preconditioner of $\boldsymbol{T}_N'\boldsymbol{T}_N$ with respect to the DCT-II, i.e.

$$\boldsymbol{M}_N := (\boldsymbol{C}_N^{II})'\delta(\boldsymbol{C}_N^{II}\boldsymbol{T}_N'\boldsymbol{T}_N(\boldsymbol{C}_N^{II})')\boldsymbol{C}_N^{II}. \tag{5.1}$$

In this section, we prove that under certain assumptions on $\boldsymbol{T}_N$, the eigenvalues of the preconditioned matrix $\boldsymbol{M}_N^{-1}\boldsymbol{T}_N'\boldsymbol{T}_N$ are clustered at 1. We follow the lines of Chan. First, we show that for all $\varepsilon > 0$ and N sufficiently large, the matrix $\boldsymbol{T}_N'\boldsymbol{T}_N - \boldsymbol{M}_N$ splits into a matrix of low rank independent of N and a matrix with spectral norm smaller than ε. Then we apply Cauchy's interlace theorem to verify that the eigenvalues of

$$\boldsymbol{M}_N^{-1}(\boldsymbol{T}_N'\boldsymbol{T}_N - \boldsymbol{M}_N) = \boldsymbol{M}_N^{-1}\boldsymbol{T}_N'\boldsymbol{T}_N - \boldsymbol{I}_N$$

are clustered at 0. Again, we drop the index N, if the dimension of the matrices follows from the context.

In preparation of Theorem 5.3, we provide the following two lemmata.

Lemma 5.1. *Let $a = \{a_k\}_{k=0}^{\infty} \in l_1$ and $b = \{b_k\}_{k=0}^{\infty} \in l_1$. Then, for all $\varepsilon > 0$, there exists $m = m(\varepsilon)$ such that for all $N \geqslant 2m$ the Hankel matrix*

$$\boldsymbol{H} := \text{hankel}((a_0, a_1, \ldots, a_{N-1} + b_{N-1}), (b_0, b_1, \ldots, a_{N-1} + b_{N-1}))$$

splits as $\boldsymbol{H} = \boldsymbol{V}_H + \boldsymbol{W}_H$, where

$$\boldsymbol{W}_H := \text{hankel}((a_0, \ldots, a_{m-1}, \boldsymbol{o}_{N-m}), (b_0, \ldots, b_{m-1}, \boldsymbol{o}_{N-m}))$$

is a matrix of rank $\leqslant 2m$ and where $\boldsymbol{V}_H := \boldsymbol{H} - \boldsymbol{W}_H$ satisfies $||\boldsymbol{V}_H||_2 < \varepsilon$.

Proof. Since $a, b \in l_1$, there exists for all $\varepsilon > 0$ an integer $m = m(\varepsilon)$ such that

$$\sum_{k=m}^{\infty} |a_k| < \varepsilon/2, \qquad \sum_{k=m}^{\infty} |b_k| < \varepsilon/2.$$

Now the assertion follows from $||\boldsymbol{V}_H||_2 \leqslant ||\boldsymbol{V}_H||_1 < \varepsilon/2 + \varepsilon/2$. □

Lemma 5.2. *Let $t = \{t_k\}_{k\in\mathbb{Z}} \in l_1$ with $||t||_1 = \tau$ be given. Further let $\boldsymbol{T} = \boldsymbol{T}_N$ and $\boldsymbol{H} = \boldsymbol{H}_N$ be the corresponding Nth Toeplitz matrix and Nth Hankel matrix, respectively. Then, for all $\varepsilon > 0$, there exists $m = m(\varepsilon)$ such that for all $N \geqslant 4m$*

$$\boldsymbol{HT} + \boldsymbol{T}'\boldsymbol{H} + \boldsymbol{H}^2 = \boldsymbol{V} + \boldsymbol{W}, \tag{5.2}$$

where $||\boldsymbol{V}||_2 < \varepsilon$ and

$$\boldsymbol{W} := \left(w_{j,k}\right)_{j,k=0}^{N-1} \quad \text{with} \quad w_{j,k} = 0 \quad \text{for } 2m \leqslant j + k \leqslant 2N - 2 - 2m. \tag{5.3}$$

Proof. By construction, we have $||\boldsymbol{T}||_2 \leqslant \tau$, $||\boldsymbol{H}||_2 \leqslant \tau$. Since $t \in l_1$, there exists $m = m(\varepsilon)$ such that

$$\sum_{|k|=m+1}^{\infty} |t_k| < \frac{\varepsilon}{6\tau}. \tag{5.4}$$

Then

$$\boldsymbol{T} = \boldsymbol{T}_\varepsilon + \boldsymbol{T}_B, \qquad \boldsymbol{H} = \boldsymbol{H}_\varepsilon + \boldsymbol{H}_B \tag{5.5}$$

with

$$\boldsymbol{T}_\varepsilon := \text{toeplitz}((\boldsymbol{0}_{m+1}, t_{-(m+1)}, \ldots, t_{-(N-1)}), (\boldsymbol{o}_{m+1}, t_{m+1}, \ldots, t_{N-1})),$$

$$\boldsymbol{T}_B := \text{toeplitz}((t_0, \ldots, t_{-m}, \boldsymbol{o}_{N-m-1}), (t_0, \ldots, t_m, \boldsymbol{o}_{N-m-1})),$$

$$\boldsymbol{H}_\varepsilon := \text{hankel}((\boldsymbol{o}_m, t_{-(m+1)}, \ldots, t_{-(N-2)}, 0), (\boldsymbol{o}_m, t_{m+1}, \ldots, t_{N-2}, 0)),$$

$$\boldsymbol{H}_B := \text{hankel}((t_{-1}, \ldots, t_{-m}, \boldsymbol{o}_{N-m}), (t_1, \ldots, t_m, \boldsymbol{o}_{N-m})),$$

where we obtain by Eq. (5.4) that $||\boldsymbol{T}_\varepsilon||_2 \leqslant \varepsilon/6\tau$, $||\boldsymbol{H}_\varepsilon||_2 \leqslant \varepsilon/6\tau$. Substituting Eq. (5.5) into Eq. (5.2), we obtain the desired decomposition

$$\begin{aligned}HT + T'H + H^2 &= (H_\varepsilon + H_B)(T_\varepsilon + T_B) \\ &\quad + (T'_\varepsilon + T'_B)(H_\varepsilon + H_B) + (H_\varepsilon + H_B)^2 \\ &= (HT_\varepsilon + T'_\varepsilon H + H_\varepsilon T_B + T'_B H_\varepsilon + H_\varepsilon H + H_B H_\varepsilon) \\ &\quad + (T'_B H_B + H_B T_B + H_B^2). \qquad \square\end{aligned}$$

Theorem 5.3. *Let $t = \{t_k\}_{k\in\mathbb{Z}} \in l_1$ with $\|t\|_1 = \tau$ be given and let $\boldsymbol{T} = \boldsymbol{T}_N$ be the corresponding Nth Toeplitz matrix. Then, for all $\varepsilon > 0$, there exists $m = m(\varepsilon)$ such that for all $N \geqslant 2m$*

$$T'T = C'DC + V + W,$$

where $\boldsymbol{D}$ denotes some diagonal matrix,

$$W := \left(w_{j,k}\right)_{j,k=0}^{N-1} \text{ with } w_{j,k} = 0 \quad \text{for } m \leqslant j + k \leqslant 2N - 2 - m,$$

and where $\|V\|_2 < \varepsilon$.

Proof. Let $\boldsymbol{H} = \boldsymbol{H}_N$ denote the Nth Hankel matrix associated with t. Then it follows by Theorem 3.1(ii) that

$$\begin{aligned}T'T &= (T' + H - H)(T + H - H) \\ &= (C'D_aC + S'\tilde{D}_bC - H)(C'D_aC + C'\tilde{D}_bS - H),\end{aligned}$$

where

$$\begin{aligned}a = a_{N+1} &:= (2t_0, t_{-1} + t_1, \ldots, t_{-(N-1)} + t_{N-1}, 0)' \in \mathbb{R}^{N+1}, \\ b = b_{N-1} &:= (t_{-1} - t_1, \ldots, t_{-(N-1)} - t_{N-1})' \in \mathbb{R}^{N-1}\end{aligned}$$

and

$$\begin{aligned}D_a &= \operatorname{diag}(d_0, \ldots, d_{N-1}, d_N)', \qquad (d_0, \ldots, d_N)' := \tilde{C}^I_{N+1} a, \\ \tilde{D}_b &= \operatorname{diag}(0, \tilde{d}_1, \ldots, \tilde{d}_{N-1}, 0)', \qquad (\tilde{d}_0, \ldots, d_{N-1})' := \tilde{S}^I_{N-1} b.\end{aligned}$$

By $C^{II}_N(C^{II}_N)' = I_N$, we further obtain that

$$\begin{aligned}T'T &= C'D_a^2C + S'\tilde{D}_b^2S + C'D_a\tilde{D}_bS + S'\tilde{D}_bD_aC \\ &\quad - H(T + H) - (T' + H)H + H^2\end{aligned}$$

and by Theorem 3.1(ii) that

$$T'T = C'D_a^2C + C'\tilde{D}_b^2C - H_{\tilde{D}_b^2} + \tilde{H}_{\tilde{D}_bD_a} - (HT + T'H + H^2) \tag{5.6}$$

with

$$\boldsymbol{H}_{\tilde{D}_b^2} := \operatorname{shankel}(u_1, \ldots, u_N),$$
$$(u_0, \ldots, u_N)' := \frac{2}{N} \tilde{\boldsymbol{C}}^I_{N+1} (0, \tilde{d}_1^2, \ldots, \tilde{d}_{N-1}^2, 0)',$$
$$\tilde{\boldsymbol{H}}_{\tilde{D}_b D_a} := \operatorname{ahankel}(v_1, \ldots, v_{N-1}, 0),$$
$$(v_1, \ldots, v_{N-1})' := \frac{2}{N} \tilde{\boldsymbol{S}}^I_{N-1} (d_1 \tilde{d}_1, \ldots, d_{N-1} \tilde{d}_{N-1})'.$$

Set

$$\hat{\boldsymbol{H}} := \boldsymbol{H}_{\tilde{D}_b^2} - \tilde{\boldsymbol{H}}_{\tilde{D}_b D_a} = \operatorname{hankel}((w_1, \ldots, w_N), (w_{-1}, \ldots, w_{-N})) \tag{5.7}$$

with $w_N = w_{-N} = u_N$, $w_k = u_k - v_k$, $w_{-k} = u_k + v_k$ $(k = 1, \ldots, N-1)$. Then we have for all $N \in \mathbb{N}$ that

$$w_N = \frac{2}{N} \sum_{k=1}^{N-1} (-1)^k \tilde{d}_k^2,$$
$$|w_N| \leqslant \frac{2}{N} ||\tilde{\boldsymbol{S}}^I_{N-1} \boldsymbol{b}||_2^2 \leqslant ||\boldsymbol{b}||_2^2 \leqslant ||\boldsymbol{b}||_1^2 \leqslant \tau^2. \tag{5.8}$$

Moreover, we get by Theorem 3.1(i) for the first row $\boldsymbol{w} := (w_1, \ldots, w_{N-1})'$ of $\hat{\boldsymbol{H}}$ that

$$\begin{aligned} \boldsymbol{w} = {} & \frac{2}{N} \boldsymbol{R}'_N \tilde{\boldsymbol{C}}^{\,I}_{N+1} \operatorname{diag}(0, \tilde{d}_1, \ldots, \tilde{d}_{N-1}, 0)' \boldsymbol{R}_N \tilde{\boldsymbol{S}}^I_{N-1} \boldsymbol{b} \\ & - \frac{2}{N} \tilde{\boldsymbol{S}}^{\,I}_{N-1} \boldsymbol{R}'_N \operatorname{diag}(d_0, \ldots, d_N)' \boldsymbol{R}_N \tilde{\boldsymbol{S}}^I_{N-1} \boldsymbol{b} \\ = {} & \operatorname{toeplitz}\big((-t_0, -t_1, \ldots, -t_{N-2}), (-t_0, -t_{-1}, \ldots, -t_{-(N-2)})\big)\boldsymbol{b} \\ & + \operatorname{hankel}\big((t_{-2}, \ldots, t_{-(N-2)}, b_{N-1}, 0), (t_2, \ldots, t_{N-2}, -b_{N-1}, 0)\big)\boldsymbol{b}. \end{aligned}$$

Thus, we obtain for all $N \in \mathbb{N}$ that

$$||\boldsymbol{w}||_1 \leqslant 2\tau^2.$$

Similarly, we conclude that

$$||(w_{-1}, \ldots, w_{-(N-1)})'||_1 \leqslant \tau^2.$$

Together with Eq. (5.8), we see that $\hat{\boldsymbol{H}}$ satisfies the assertion of Lemma 5.1. Hence, for fixed $\varepsilon > 0$, there exists $\hat{m} = \hat{m}(\varepsilon) > 0$ such that for all $N \geqslant 2\hat{m}$

$$\hat{\boldsymbol{H}} = \hat{\boldsymbol{V}} + \hat{\boldsymbol{W}} \tag{5.9}$$

with $||\hat{\boldsymbol{V}}||_2 \leqslant \varepsilon/2$, $\hat{\boldsymbol{W}} = \operatorname{hankel}((w_1, \ldots, w_m, \boldsymbol{o}_{N-m}), (w_{-1}, \ldots, w_{-m}, \boldsymbol{o}_{N-m}))$. Furthermore, by Lemma 5.2, there exists $\tilde{m} = \tilde{m}(\varepsilon) > 0$ such that for N sufficiently large

$$(\boldsymbol{H}\boldsymbol{T} + \boldsymbol{T}'\boldsymbol{H} + \boldsymbol{H}^2) = \tilde{\boldsymbol{V}} + \tilde{\boldsymbol{W}} \tag{5.10}$$

with $\|\tilde{\boldsymbol{V}}\|_2 \leqslant \varepsilon/2$ and with a low rank matrix $\tilde{\boldsymbol{W}}$ of the form Eq. (5.3). Applying Eqs. (5.9) and (5.10) in Eq. (5.6), we obtain the assertion

$$\boldsymbol{T}'\boldsymbol{T} = \boldsymbol{C}'(\boldsymbol{D}_a^2 + \tilde{\boldsymbol{D}}_b^2)\boldsymbol{C} + (\tilde{\boldsymbol{V}} - \hat{\boldsymbol{V}}) + \tilde{\boldsymbol{W}} - \hat{\boldsymbol{W}}$$

with $m := \max\{\hat{m}, 2\tilde{m}\}$. □

Lemma 5.4. *For $m > 0$ and $N > 2m$, let $\boldsymbol{V} \in \mathbb{R}^{N,N}$ with $\|\boldsymbol{V}\|_2 < \varepsilon/2$ and*

$$\boldsymbol{W} := \left(w_{j,k}\right)_{j,k=0}^{N-1} \text{ with } w_{j,k} = 0 \quad \text{for } m \leqslant j + k \leqslant 2N - 2 - m,$$

be given. Set $\omega := \sum_{j,k=0}^{N-1} |w_{j,k}|$. Then, for $N > 4\omega/\varepsilon$,

$$\|\delta(\boldsymbol{C}_N^{II}(\boldsymbol{V} + \boldsymbol{W})(\boldsymbol{C}_N^{II})')\|_2 < \varepsilon.$$

Note that for fixed m, the value ω does not depend on N.

Proof. On the one hand, we obtain that

$$\|\delta(\boldsymbol{C}_N^{II}\boldsymbol{V}(\boldsymbol{C}_N^{II})')\|_2 \leqslant \|\boldsymbol{C}_N^{II}\boldsymbol{V}(\boldsymbol{C}_N^{II})'\|_2 = \|\boldsymbol{V}\|_2 < \varepsilon/2,$$

and on the other hand that

$$\begin{aligned}
&\|\delta(\boldsymbol{C}_N^{II}\boldsymbol{W}(\boldsymbol{C}_N^{II})')\|_2 \\
&\leqslant \max_{n=0,\ldots,N-1} \left| \frac{2}{N}(\varepsilon_n^N)^2 \sum_{j,k=0}^{N-1} w_{j,k} \cos\frac{n(2j+1)\pi}{2N} \cos\frac{n(2k+1)\pi}{2N} \right| \\
&\leqslant 2\omega/N < \varepsilon/2 \quad (N > 4\omega/\varepsilon).
\end{aligned}$$

Now summation implies the assertion. □

Theorem 5.5. *Let $t = \{t_k\}_{k\in\mathbb{Z}} \in l_1$ with $\|t\|_1 = \tau$ be given and let $\boldsymbol{T}_N$ be the corresponding Nth Toeplitz matrix. Moreover, assume that the singular values of $\boldsymbol{T}_N$ are larger than $\gamma > 0$ for all $N \in \mathbb{N}$. Let $\boldsymbol{M}_N = \boldsymbol{M}_N(\boldsymbol{T}_N'\boldsymbol{T}_N)$ denote the optimal preconditioner of $\boldsymbol{T}_N'\boldsymbol{T}_N$ in $\mathscr{A}_{C_N^{II}}$. Then the eigenvalues of $\boldsymbol{M}_N^{-1}\boldsymbol{T}_N'\boldsymbol{T}_N$ are clustered at 1.*

Proof. By Corollary 4.2, we see that $\|\boldsymbol{M}_N^{-1}\|_2 < 1/\gamma$ for all $N \in \mathbb{N}$. Let $\varepsilon > 0$ be fixed. Then we obtain by Eq. (5.1), Theorem 5.3 and Lemma 5.4 that for N sufficiently large, there exists $M \in \mathbb{N}$ independent of N such that

$$\begin{aligned} T_N' T_N - M_N &= V_N + W_N - (C_N^{II})' \delta(C_N^{II}(V_N + W_N)(C_N^{II})') C_N^{II} \\ &= U_N + W_N, \end{aligned}$$

where W_N is of low rank M and where $\|U_N\|_2 < \varepsilon\gamma$. Now

$$\begin{aligned} M_N^{-1/2} T_N' T_N M_N^{-1/2} - I_N &= M_N^{-1/2} U_N M_N^{-1/2} + M_N^{-1/2} \tilde{W}_N M_N^{-1/2} \\ &= \tilde{V}_N + \tilde{W}_N, \end{aligned}$$

where $\|\tilde{V}_N\|_2 \leqslant \varepsilon$ and where the rank of $\tilde{W}_N$ is at most M. Since $\tilde{V}_N$ and $\tilde{W}_N$ are symmetric matrices, we can apply Cauchy's interlace theorem [37], which implies that for N sufficiently large, at most M eigenvalues of $\tilde{V}_N + \tilde{W}_N$ have absolute value greater than ε. Now the assertion follows since $M_N^{-1/2} T_N' T_N M_N^{-1/2}$ and $M_N^{-1} T_N' T_N$ possess the same eigenvalues. □

Remark. Under the above assumptions on T_N it was proved that the eigenvalues of $(M_N^{-1} T_N)^*(M_N^{-1} T_N)$, where $M_N = C_N(T_N)$ denotes the optimal circulant preconditioner of T_N, are clustered at 1 [13,34]. In general, the eigenvalues of $(M_N^{-1} T_N)'(M_N^{-1} T_N)$ are not clustered at 1, if M_N is the optimal trigonometric preconditioner of T_N. If $T_N = (t_{j-k})_{j,k=0}^{N-1}$ with $t_0 = 1$ and $t_k = -t_{-k}$ $(k = 1, \ldots, N-1)$, then the optimal trigonometric preconditioners of T_N are $M_N = I_N$, i.e. we have no preconditioning. If, for example, $t_{-1} = -t_1 = 2$ and $t_k = 0$ $(|k| > 1)$, then the matrices T_{2N+1} $(N \in \mathbb{N})$ satisfy the assumption of Theorem 5.5. However, the eigenvalues of $T_{2N+1}' T_{2N+1}$ are given by $9 - 8\cos(j\pi/(N+1))$ $(j = 0, \ldots, N)$. □

6. Construction of optimal preconditioners of $T'T$

In this section, we explain how optimal trigonometric preconditoners of $T_N' T_N$ can be constructed *with* $\mathrm{O}(N \log N)$ *arithmetical operations*. In contrast to the construction of optimal trigonometric preconditioners of T_N, we are confronted with the fact that $T_N' T_N$ is not a Toeplitz matrix. Again, we consider only C_N^{II}-preconditioners. The approach for the DST-I, DST-II, DST-IV and DCT-IV follows the same lines. For the construction of the optimal preconditioner $M_N = M_N(T_N' T_N)$ we use the representation Eq. (4.7) of M_N with the basis $\{B_k^{II} : k = 0, \ldots, N-1\}$ of $\mathcal{A}_{C_N^{II}}$ [6,17]:

$$B_k^{II} := (C_N^{II})' \operatorname{diag}(U_k(c_l))_{l=0}^{N-1} C_N^{II},$$

where $c_l := \cos l\pi/N$ and where U_k denotes the kth *Chebyshev polynomial of second kind*

$$U_k(x) := \sin((k+1)\arccos x)/\sin(\arccos x) \quad (x \in (-1, 1)).$$

Moreover, we apply that $\{B_k^I : k = 0, \ldots, N-1\}$ with

$$\begin{aligned} B_k^I &:= (C_N^{II})' \operatorname{diag}(T_k(c_l))_{l=0}^{N-1}\, C_N^{II} \\ &= \text{stoeplitz}\, e_k' + \text{shankel}\, e_{k-1}' \quad (k = 1, \ldots, N-1), \end{aligned} \tag{6.1}$$

$e_{-1} := o_N$, and with the kth *Chebyshev polynomial of first kind*

$$T_k(x) := \cos(k \arccos x) \quad (x \in [-1, 1]),$$

is another basis of $\mathcal{A}_{C_N^{II}}$. Both bases are related by

$$\begin{aligned} B_0^I &= B_0^{II} = I, \qquad B_1^{II} = 2B_1^I, \\ B_j^{II} &= B_{j-2}^{II} + 2B_j^I \quad (j = 2, \ldots, N-1), \end{aligned} \tag{6.2}$$

where the last equation follows by $U_j = U_{j-2} + 2T_j$. Now we have by Eq. (4.7) that

$$M_N = \sum_{k=0}^{N-1} \alpha_k\, B_k^{II}$$

with

$$\boldsymbol{\alpha} = G^{-1}\boldsymbol{\beta}^{II}, \qquad G := \left(\langle B_j^{II}, B_k^{II} \rangle\right)_{j,k=0}^{N-1}, \qquad \boldsymbol{\beta}^{II} := \left(\langle T_N' T_N, B_j^{II} \rangle\right)_{j=0}^{N-1}.$$

Clearly, we are not interested in M_N itself, but in the diagonal matrix diag $\boldsymbol{d}$ with

$$M_N = (C_N^{II})' \,(\operatorname{diag} \boldsymbol{d})\, C_N^{II}. \tag{6.3}$$

If $\boldsymbol{\alpha}$ is known, then we obtain diag $\boldsymbol{d}$ by

$$\operatorname{diag} \boldsymbol{d} = \sum_{k=0}^{N-1} \alpha_k C_N^{II} B_k^{II} (C_N^{II})' = \sum_{k=0}^{N-1} \alpha_k \operatorname{diag}\, (U_k(c_l))_{l=0}^{N-1},$$

i.e. by definition of U_k by

$$d_0 = \sum_{k=0}^{N-1} (k+1)\, \alpha_k, \qquad d_k = \frac{\hat{\alpha}_k}{\sin k\pi/N} \quad (k = 1, \ldots, N-1), \tag{6.4}$$

$$(\hat{\alpha}_k)_{k=1}^{N-1} = \tilde{S}_{N-1}^I\, (\alpha_{k-1})_{k=1}^{N-1}.$$

Thus the construction of $\boldsymbol{d}$ from given $\boldsymbol{\alpha}$ requires $O(N \log N)$ arithmetical operations. It remains to find an efficient construction of the coefficient vector $\boldsymbol{\alpha}$. Therefore we use the following lemmata.

Lemma 6.1. *The inverse matrix of* $\boldsymbol{G} := \left(\langle \boldsymbol{B}_j^{II}, \boldsymbol{B}_k^{II} \rangle\right)_{j,k=0}^{N-1}$, *is given by*

$$\boldsymbol{G}^{-1} = \frac{1}{2N} \begin{pmatrix} 3 & 0 & -1 & & & & \\ 0 & 2 & 0 & -1 & & & \\ -1 & 0 & 2 & 0 & -1 & & \\ & \ddots & \ddots & \ddots & \ddots & & \\ & & -1 & 0 & 2 & 0 & -1 \\ & & & -1 & 0 & 3 & -2 \\ & & & & -1 & -2 & \frac{3N-2}{N} \end{pmatrix},$$

such that the vector multiplication with $\boldsymbol{G}^{-1}$ *can be computed with* $\mathrm{O}(N)$ *arithmetical operations.*

Proof. We show that $\boldsymbol{G}^{-1}\,\boldsymbol{G} = \boldsymbol{I}$. By definition, we have for the (k,l)th entry of $\boldsymbol{G}$ that

$$\begin{aligned} g_{k,l} &= \langle \boldsymbol{B}_k^{II}, \boldsymbol{B}_l^{II} \rangle = \operatorname{tr}\left((\boldsymbol{B}_l^{II})'\, \boldsymbol{B}_k^{II}\right) \\ &= \sum_{j=0}^{N-1} U_l(c_j)\, U_k(c_j). \end{aligned}$$

Then we get for $l = 2, \ldots, N-3$ that

$$2g_{k,l} - g_{k,l-2} - g_{k,l+2} = \sum_{j=0}^{N-1} U_k(c_j)\, \left(2U_l(c_j) - U_{l-2}(c_j) - U_{l+2}(c_j)\right)$$

and further since

$$2U_l(x) - U_{l-2}(x) - U_{l+2}(x) = 4(1-x^2)U_l(x)$$

and by definition of U_l that

$$\begin{aligned} 2g_{k,l} - g_{k,l-2} - g_{k,l+2} &= 4\sum_{j=1}^{N-1} \sin\frac{(k+1)j\pi}{N} \sin\frac{(l+1)j\pi}{N} \\ &= 2N\,\delta_{k,l} \end{aligned}$$

$(k = 0, \ldots, N-1; l = 2, \ldots, N-3)$. Straightforward computation for $l = 0, 1, N-2, N-1$ completes the proof. □

Set

$$\boldsymbol{\beta}^I := \left(\langle \boldsymbol{T}_N' \boldsymbol{T}_N, \boldsymbol{B}_j^I \rangle\right)_{j=0}^{N-1}.$$

Then we obtain by the recurrence relation (6.2) that

$$\beta_0^{II} = \beta_0^I, \qquad \beta_1^{II} = 2\beta_1^I, \qquad \beta_k^{II} = \beta_{k-2}^{II} + \beta_k^I \quad (k = 2, \ldots, N-1). \tag{6.5}$$

Thus we can compute $\boldsymbol{\beta}''$ from $\boldsymbol{\beta}'$ with O(N) additions. The following construction of $\boldsymbol{\beta}'$ is based on an idea of Tyrtyshnikov [33]. We split $\boldsymbol{T} = \boldsymbol{T}_N = (t_{j-k})_{j,k=0}^{N-1}$ into a lower and an upper triangular Toeplitz matrix

$$\boldsymbol{T} = \boldsymbol{T}_L + \boldsymbol{T}_R$$

with diagonal entries $t_0/2$. Then we obtain that

$$\begin{aligned}\beta'_k &= \langle \boldsymbol{T}'_L \boldsymbol{T}_L, \boldsymbol{B}'_k\rangle + \langle \boldsymbol{T}'_R \boldsymbol{T}_R, \boldsymbol{B}'_k\rangle \\ &\quad + \langle \boldsymbol{T}'_L \boldsymbol{T}_R + \boldsymbol{T}'_R \boldsymbol{T}_L, \boldsymbol{B}'_k\rangle \quad (k = 0, \ldots, N-1). \end{aligned} \tag{6.6}$$

We consider the summands on the right-hand side. The matrix $\boldsymbol{T}'_L \boldsymbol{T}_R + \boldsymbol{T}'_R \boldsymbol{T}_L$ is a symmetric Toeplitz matrix. The matrices $\boldsymbol{T}'_L \boldsymbol{T}_L$ and $\boldsymbol{T}'_R \boldsymbol{T}_R$ have lost their Toeplitz structure. For $\boldsymbol{A} \in \mathbb{R}^{N,N}$, we introduce the vectors $\boldsymbol{s}(\boldsymbol{A}) = (s_k(\boldsymbol{A}))_{k=0}^{N-1}$, $\boldsymbol{h}(\boldsymbol{A}) = (h_k(\boldsymbol{A}))_{k=0}^{N-1}$ and $\tilde{\boldsymbol{h}}(\boldsymbol{A}) = (\tilde{h}_k(\boldsymbol{A}))_{k=0}^{N-1}$ by

$$\begin{aligned} s_k(\boldsymbol{A}) &:= \langle \boldsymbol{A}, \text{stoeplitz}((\varepsilon_k^N)^{-2}\boldsymbol{e}_k)\rangle \quad (k = 0, \ldots, N-1), \\ h_k(\boldsymbol{A}) &:= \langle \boldsymbol{A}, \text{hankel}(\boldsymbol{e}_k, 0)\rangle \quad (k = 0, \ldots, N-2), \\ \tilde{h}_k(\boldsymbol{A}) &:= \langle \boldsymbol{A}, \text{hankel}(0, \boldsymbol{e}_k)\rangle \quad (k = 0, \ldots, N-2). \end{aligned}$$

Set $h_{-1} := 0$ and $\tilde{h}_{-1} := 0$. Then it follows by Eq. (6.1) that

$$\langle \boldsymbol{A}, \boldsymbol{B}'_k\rangle = (\varepsilon_k^N)^2 s_k(\boldsymbol{A}) + h_{k-1}(\boldsymbol{A}) + \tilde{h}_{k-1}(\boldsymbol{A}) \quad (k = 0, \ldots, N-1). \tag{6.7}$$

Lemma 6.2. *Let* $\boldsymbol{A} := \boldsymbol{T}'_L \boldsymbol{T}_R + \boldsymbol{T}'_R \boldsymbol{T}_L$ *and let* $\boldsymbol{r} := \boldsymbol{T}'_R(t_0, t_1, \ldots, t_{N-1})'$. *Then*

$$\begin{aligned} s_k(\boldsymbol{A}) &= 2(N-k) r_k \quad (k = 0, \ldots, N-1), \\ h_0(\boldsymbol{A}) &= r_0, \qquad h_1(\boldsymbol{A}) = 2\, r_1, \\ h_k(\boldsymbol{A}) &= 2\, r_k + h_{k-2}(\boldsymbol{A}) \quad (k = 2, \ldots, N-2), \\ \tilde{h}_k(\boldsymbol{A}) &= h_k(\boldsymbol{A}) \quad (k = 0, \ldots, N-2). \end{aligned}$$

Proof. Since $\boldsymbol{A} = \text{stoeplitz}\ \boldsymbol{r}$, the assertion follows by definition of $\boldsymbol{s}, \boldsymbol{h}$ and $\tilde{\boldsymbol{h}}$. □

Lemma 6.3. *Let* $\boldsymbol{A} := \boldsymbol{T}'_R \boldsymbol{T}_R$ *and let* $\boldsymbol{r} := \left((\varepsilon_k^N)^2 t_{-k}\right)_{k=0}^{N-1}$. *Then*

$$\boldsymbol{s}(\boldsymbol{A}) = 2\, \text{hankel}(N r_0, (N-1) r_1, \ldots, r_{N-1}), (0, \ldots, 0, r_{N-1}))\boldsymbol{r}, \tag{6.8}$$

$$h_0(\boldsymbol{A}) = x_0, \qquad h_1(\boldsymbol{A}) = x_1,$$

$$h_k(\boldsymbol{A}) = h_{k-2}(\boldsymbol{A}) + x_k \quad (k = 2, \ldots, N-2), \tag{6.9}$$

$$\tilde{h}_0(\boldsymbol{A}) = y_0/2, \qquad \tilde{h}_1(\boldsymbol{A}) = y_1,$$

$$\tilde{h}_k(\boldsymbol{A}) = \tilde{h}_{k-2}(\boldsymbol{A}) + y_k \quad (k = 2, \ldots, N-2), \tag{6.10}$$

where

$$x := T_R' r,$$
$$y := \text{hankel}(2(r_0, r_1, \ldots, r_{N-1}), (-r_2, -r_3, \ldots, -r_{N-1}, 0, 2r_{N-1}))r.$$

Proof. Since T_R is an upper triangular Toeplitz matrix, we have that

$$a_{j,k} = a_{j-1,k-1} + r_j r_k \quad (j, k = 1, \ldots, N-1). \tag{6.11}$$

Consequently, we obtain that

$$s_k := 2 \sum_{j=0}^{N-k-1} a_{j+k,j} = \sum_{j=0}^{N-1-k} (N-j-k) r_{j+k} r_j$$

which yields Eq. (6.8). The recursions Eqs. (6.9) and (6.10) follow by straightforward calculation from Eq. (6.11). □

Lemma 6.4. *Let* $B := T_R' T_R$ *and let* $J :=$ *shankel* e_{N-1} *denote the* N*th counteridentity. Then*

$$s_k(B) = s_k(J\,BJ) \quad (k = 0, \ldots, N-1),$$

$$h_k(B) = \tilde{h}_k(J\,BJ) \quad (k = 0, \ldots, N-2),$$

$$\tilde{h}_k(B) = h_k(J\,BJ) \quad (k = 0, \ldots, N-2),$$

such that $s(B), h(B)$ *and* $\tilde{h}(B)$ *can be computed by Eqs.* (6.8)–(6.10).

Proof. The relations for $s(B), h(B)$ and $\tilde{h}(B)$ follow by definition of J. By

$$J\,T_L'\,T_L\,J = (J\,T_L\,J)'\,(J\,T_L\,J)$$

and since

$$J\,T_L\,J = \text{toeplitz}((t_0, \ldots, t_{N-1}), (t_0, 0, \ldots, 0))$$

is an upper triangular Toeplitz matrix, we can calculate $s(B), h(B)$ and $\tilde{h}(B)$ by Eqs. (6.8)–(6.10) with $A = JBJ$ and $r = \left((\varepsilon_k^N)^2 t_k\right)_{k=0}^{N-1}$. □

Theorem 6.5. *Let* $T_N := (t_{j-k})_{j,k=0}^{N-1}$. *Then the optimal preconditioner* $M_N \in \mathscr{A}_{C_N^{II}}$ *of* $T_N' T_N$ *can be constructed with* $O(N \log N)$ *arithmetical operations.*

Proof. We compute β^I by Eqs. (6.6) and (6.7) and by the Lemmata 6.2–6.4. Taking into account that the multiplication of a vector with a Toeplitz matrix or a Hankel matrix requires $O(N \log N)$ operations, the whole construction of β^I takes $O(N \log N)$ arithmetical operations. From β^I we compute β^{II} by

Eq. (6.5) with $O(N)$ additions. Using Lemma 6.1, we get $\boldsymbol{\alpha} := \boldsymbol{G}^{-1}\boldsymbol{\beta}^{II}$ at the cost of $O(N)$ arithmetical operations. Finally, the DST-I in Eq. (6.4) to obtain $\boldsymbol{d}$ from $\boldsymbol{\alpha}$ needs $O(N \log N)$ arithmetical operations and we are done with an arithmetical complexity of $O(N \log N)$. □

Remark. We can use similar ideas for the construction of the optimal trigonometric preconditioners with respect to the $\boldsymbol{C}_N^{IV}$, $\boldsymbol{S}_{N-1}^{I}$, $\boldsymbol{S}_N^{II}$ and $\boldsymbol{S}_N^{IV}$. For the corresponding bases of $\mathscr{A}_{O_N}$

$$\begin{aligned} \boldsymbol{B}_k^I &:= \boldsymbol{O}'_{\text{dim}} \operatorname{diag}(T_k(c_l))_{l=0}^{\text{dim}-1} \boldsymbol{O}_{\text{dim}} \\ &= \operatorname{stoeplitz} \boldsymbol{e}'_k + \operatorname{hankel}(\boldsymbol{u}'_k, \boldsymbol{v}'_k), \\ \boldsymbol{B}_k^{II} &:= \boldsymbol{O}'_{\text{dim}} \operatorname{diag}(U_k(c_l))_{l=0}^{\text{dim}-1} \boldsymbol{O}_{\text{dim}} \end{aligned}$$

we obtain

$$\boldsymbol{G}^{-1} = \frac{1}{K} \begin{pmatrix} 3 & 0 & -1 & & & & & \\ 0 & 2 & 0 & -1 & & & & \\ -1 & 0 & 2 & 0 & -1 & & & \\ & \ddots & \ddots & \ddots & \ddots & & & \\ & & -1 & 0 & 2 & 0 & -1 \\ & & & -1 & 0 & g_1 & g_2 \\ & & & & -1 & g_2 & g_3 \end{pmatrix},$$

where

$\boldsymbol{O}_{\text{dim}}$	c_l	dim	$\boldsymbol{u}_k$	$\boldsymbol{v}_k$	K	g_1	g_2	g_3
$\boldsymbol{C}_N^{II}$	$\cos \frac{l\pi}{N}$	N	$\boldsymbol{e}_{k-1}$	$\boldsymbol{e}_{k-1}$	$2N$	3	-2	$\frac{3N-2}{N}$
$\boldsymbol{C}_N^{IV}$	$\cos \frac{(2l+1)\pi}{2N}$	N	$\boldsymbol{e}_{l-1}$	$-\boldsymbol{e}_{k-1}$	$2N$	1	0	1
$\boldsymbol{S}_{N-1}^{I}$	$\cos \frac{(l+1)\pi}{N}$	$N-1$	$-\boldsymbol{e}_{k-2}$	$-\boldsymbol{e}_{k-2}$	$2N+2$	2	0	3
$\boldsymbol{S}_N^{II}$	$\cos \frac{(l+1)\pi}{N}$	N	$-\boldsymbol{e}_{k-1}$	$-\boldsymbol{e}_{k-1}$	$2N$	3	2	$\frac{3N-2}{N}$
$\boldsymbol{S}_N^{IV}$	$\cos \frac{(2l+1)\pi}{2N}$	N	$-\boldsymbol{e}_{k-1}$	$\boldsymbol{e}_{k-1}$	$2N$	1	0	1

Note that the construction of the optimal preconditioner with respect to $\boldsymbol{C}_N^{IV}$ or $\boldsymbol{S}_N^{IV}$ is especially simple.

7. Numerical results

Finally, we present examples of nonsymmetric Toeplitz systems (1.1) for which the preconditioning of the normal equation by an optimal trigonometric preconditioner $M_N = M_N(T'_N T_N)$ of $T'_N T_N$ significantly accelerates the convergence of the CG-method. The algorithms were realized for the optimal preconditioners with respect to the DCT-II, the DST-II, the DCT-IV and the DST-IV,respectively. Clearly, we can also use the DST-I preconditioner. Note that for *symmetric* Toeplitz matrices, the DST-I preconditioned CG-method for the solution of Eq. (1.1) shows a similar convergence behaviour as the DST-II preconditioned CG-method.

The fast computation of the preconditioners in the initial step and the computation of the preconditioned CG-method (PCG-method) were implemented in MATLAB and tested on a Sun SPARCstation 20. The fast trigonometric transforms appearing both in the initialization and in the PCG-steps were taken from the C-implementation based on [2,31] by using the cmex-program.

As transform length we choose $N = 2^n$. The right-hand side b_N of Eq. (1.1) is the vector consisting of N ones. The PCG-method starts with the zero vector and stops if $\|r^{(j)}\|_2/\|r^{(0)}\|_2 < 10^{-7}$, where $r^{(j)}$ denotes the residual vector after j iterations. Our test matrices are the following four Toeplitz matrices $T_N = (t_{j-k})_{j,k=0}^{N-1}$:

(i) (see [26])

$$t_n = \begin{cases} 1/\log(2-n), & n \leqslant -1, \\ 1/(\log(2)+1), & n = 0, \\ 1/(1+n), & n \geqslant 1. \end{cases}$$

(ii) (see [26])

$$t_n = \begin{cases} 2, & n = 0, \\ -0.7t_{n+1}, & n \leqslant -1, \\ 0.9t_{n-1}, & n \geqslant 1. \end{cases}$$

(iii) Here we use the Toeplitz matrices T_N arising from the generating function

$$f(x) = x^2 e^{ix}.$$

(iv)

$$t_n = \begin{cases} -1.5, & n = -1, \\ 2, & n = 0, \\ 0.5, & n = 1, \\ 0, & \text{else}. \end{cases}$$

As expected, also for large transform lengths N, the initialisation and each PCG-step can be computed in a very fast way which reflects the arithmetic complexity of $O(N \log N)$ for these computations. We compare three different methods. The second columns of the following tables show the number of iterations of the PCG-method applied to the normal equation without preconditioning. The columns 3 and 4 contain the numbers of iterations required by the CG-method applied to

$$(M_N^{-1}T_N)'(M_N^{-1}T_N)x_N = (M_N^{-1}T_N)'M_N^{-1}b_N, \tag{7.1}$$

where M_N denotes the optimal trigonometric preconditioner of T_N with respect to the DCT-II and the DST-II, respectively. The columns 5–8 contain the numbers of iterations required by the PCG-method applied to

$$M_N^{-1}T_N'T_Nx_N = M_N^{-1}T_N'b_N, \tag{7.2}$$

where M_N denotes the optimal trigonometric preconditioner of $T_N'T_N$ with respect to the DCT-II, DST-II, DCT-IV and the DST-VI, respectively. We compare our results with the GMRES-method applied to the original problem

$$M_N^{-1}T_Nx_N = M_N^{-1}b_N. \tag{7.3}$$

Again, M_N denotes the optimal trigonometric or circulant preconditioner of T_N. Here we avoid the transition of Eq. (1.1) to the normal equation. The ninth columns contain the number of iteration steps of the GMRES-method without preconditioning. Since the optimal trigonometric preconditioners are symmetric, we expect no good convergence behaviour of the corresponding PCG-method. The tenth columns confirm these expectations. Finally, the last columns contain the number of iterations of the GMRES-method with the optimal circulant preconditioner (see Ref. [35]). The number of iteration steps in the last columns of the Tables 1, 2 and 4 are compatible with the number of iteration steps required by our new preconditioning method. Our preconditioning uses the normal equation of Eq. (1.1) but avoids complex arithmetic. Note that very ill-conditioned Toeplitz matrices will be handled in a forthcoming paper.

Although not all matrices in our examples fulfil the assumptions of Theorem 5.5, the preconditioning with an optimal trigonometric preconditioner of

Table 1
Number of iterations for example (i)

		CG applied to Eq. (7.1)		PCG applied to Eq. (7.2)				GMRES applied to Eq. (7.3)		
n	I_N	C_N^{II}	S_N^{II}	C_N^{II}	S_N^{II}	C_N^{IV}	S_N^{IV}	I_N	C_N^{II}	F_N
7	24	11	10	8	15	14	11	22	15	7
8	32	12	11	8	17	15	11	28	18	8
9	43	14	13	8	19	17	11	34	21	8
10	57	18	16	9	20	19	11	43	25	8
11	86	20	19	9	20	20	12	54	30	8
12	121	25	22	9	22	22	12	>70	36	8
13	176	30	28	9	22	22	12	>70	40	8

Table 2
Number of iterations for example (ii)

		CG applied to Eq. (7.1)		PCG applied to Eq. (7.2)				GMRES applied to Eq. (7.3)		
n	I_N	C_N^{II}	S_N^{II}	C_N^{II}	S_N^{II}	C_N^{IV}	S_N^{IV}	I_N	C_N^{II}	F_N
7	34	44	44	9	12	9	14	>70	>70	8
8	43	47	50	8	11	8	13	>70	>70	8
9	53	50	52	7	10	8	12	>70	>70	8
10	59	50	53	7	9	7	11	>70	>70	8
11	50	50	53	6	9	7	10	>70	>70	8
12	58	49	53	6	8	7	10	>70	>70	8
13	56	48	54	6	8	7	9	>70	>70	8

$\boldsymbol{T}_N' \boldsymbol{T}_N$ accerelates the convergence of the CG-method significantly. Further extensive numerical tests (see Ref. [28]) with matrices from [8,9] show that at least one of our trigonometric preconditioners works better than the circulant preconditioners.

Except for the second example, where the condition numbers of $\boldsymbol{T}_N$ are bounded for $N \to \infty$, the number of iterations differs, if we apply the preconditioners with respect to DCT-II and DST-II. Heuristically, this can be explained by the different structures of $\mathscr{A}_{C_N^{II}}$ and $\mathscr{A}_{S_N^{II}}$ and how "good" our example matrices fit into this structure (see Table 3). A general criterion for the choice of the optimal trigonometric preconditioner would be interesting. In this direction, it is remarkable, that the optimal preconditioner $\boldsymbol{M}_N(\boldsymbol{T}_N) \in \mathscr{A}_{C_N^{II}}$ of the auto-covariance matrix $\boldsymbol{T}_N := (\rho^{|j-k|})_{j,k=0}^{N-1}$ is "asymptotically equivalent" to $\boldsymbol{T}_N$ if $N \to \infty$, $\rho \to 1$, while the optimal preconditioner $\boldsymbol{M}_N(\boldsymbol{T}_N) \in \mathscr{A}_{S_N^{II}}$ of $\boldsymbol{T}_N$ is "asymptotically equivalent" to $\boldsymbol{T}_N$ if $N \to \infty$, $\rho \to 0$ [29].

Now let $\boldsymbol{T} := (t_{j-k})_{j=0,k=0}^{M-1,N-1}$ denote a real $(M \times N)$ – Toeplitz matrix of rank N $(M \geqslant N)$. We are interested in the solution of least squares problems

Table 3
Number of iterations for example (iii)

		CG applied to Eq. (7.1)		PCG applied to Eq. (7.2)				GMRES applied to Eq. (7.3)		
n	I_N	C_N^{II}	S_N^{II}	C_N^{II}	S_N^{II}	C_N^{IV}	S_N^{IV}	I_N	S_N^{II}	F_N
5	84	72	68	29	21	47	24	32	32	14
6	311	124	176	52	26	84	39	64	53	16
7	1226	264	412	116	33	173	63	>70	>70	18
8	5220	626	980	256	40	405	136	>70	>70	21
9	>1000	1741	3341	664	74	1031	310	>70	>70	27

Table 4
Number of iterations for example (iv)

		CG applied to Eq. (7.1)		PCG applied to Eq. (7.2)				GMRES applied to Eq. (7.3)		
n	I_N	C_N^{II}	S_N^{II}	C_N^{II}	S_N^{II}	C_N^{IV}	S_N^{IV}	I_N	S_N^{II}	F_N
6	88	21	37	21	9	25	16	64	36	11
7	201	31	67	27	8	31	19	>70	53	11
8	435	45	125	36	8	39	24	>70	>70	12
9	929	66	294	47	9	72	32	>70	>70	12

Table 5
Number of iterations for example (v)

			PCG applied to Eq. (7.2)			
n	M	I_N	C_N^{II}	S_N^{II}	C_N^{IV}	S_N^{IV}
4	32	22	9	20	19	18
5	64	46	10	29	28	24
6	128	75	9	28	25	20
7	256	125	8	23	19	17
8	512	167	8	18	16	15
9	1024	193	8	16	15	13

$$\min \|\boldsymbol{b}_N - \boldsymbol{T}\boldsymbol{x}_N\|_2 \tag{7.4}$$

by the PCG-method applied to the corresponding normal equation. Circulant preconditioning of the normal equation of Eq. (7.4) was considered in [8,9] (see Table 4). Following the lines of Section 6, the computation of the optimal trigonometric preconditioner is straightforward. We consider the following example (v): (see [9])

$$t_k := \begin{cases} \mathrm{e}^{-0.1(k-1)^2}, & k = -N+1,\ldots,0, \\ \mathrm{e}^{-0.1k^2}, & k = 1,\ldots,M-1. \end{cases}$$

The number of iteration steps for the different optimal trigonometric preconditioners are contained in Table 5.

Acknowledgements

The authors wish to thank M. Tasche for his useful hints concerning the relations between trigonometric transforms and Toeplitz-plus-Hankel matrices. Many thanks to the referees for their valuable comments, in particular for pointing out various references to us.

References

[1] O. Axelsson, Iterative Solution Methods, Cambridge University Press, Cambridge, 1996.

[2] G. Baszenski, M. Tasche, Fast polynomial multiplication and convolution related to the discrete cosine transform, Linear Algebra Appl. 252 (1997) 1–25.

[3] D. Bini, F.D. Benedetto, A new preconditioner for the parallel solution of positive definite Toeplitz systems, in: Proceedings of the Second ACM Symposium on Parallel Algorithms and Architectures, Crete, 1990, pp. 220–223.

[4] D. Bini, P. Favati, On a matrix algebra related to discrete Hartley transform, SIAM J. Matrix Anal. Appl. 14 (1993) 500–507.

[5] E. Boman, I. Koltracht, Fast transform based preconditioners for Toeplitz equations, SIAM J. Matrix Anal. Appl. 16 (1995) 628–645.

[6] E. Bozzo, C.D. Fiore, On the use of certain matrix algebras associated with discrete trigonometric transforms in matrix displacement decomposition, SIAM J. Matrix Anal. Appl. 16 (1995) 312–326.

[7] R.H. Chan, T.F. Chan, C.K. Wong, Cosine transform based preconditioners for total variation minimization problems in image processing, Technical Report 95–8, Chinese University of Hong Kong, 1995.

[8] R.H. Chan, J. Nagy, R. Plemmons, Circulant preconditioned Toeplitz least square problems, SIAM J. Matrix Anal. Appl. 15 (1994) 80–97.

[9] R.H. Chan, J. Nagy, R. Plemmons, Displacement preconditioners for Toeplitz least squares iterations, Elec. Trans. Numer. Anal. 2 (1994) 44–56.

[10] R.H. Chan, M.K. Ng, Conjugate gradient methods of Toeplitz systems, SIAM Rev. 38 (1996) 427–482.

[11] R.H. Chan, M.K. Ng, C.K. Wong, Sine transform based preconditioners for symmetric Toeplitz systems, Linear Algebra Appl. 232 (1996) 237–259.

[12] R.H. Chan, G. Strang, Toeplitz systems by conjugate gradients with circulant preconditioner, SIAM J. Sci. Statist. Comput. 10 (1989) 104–119.

[13] R.H. Chan, M.-C. Yeung, Circulant preconditioners constructed from kernels, SIAM J. Numer. Anal. 29 (1992) 1093–1103.

[14] T.F. Chan, An optimal circulant preconditioner for Toeplitz systems, SIAM J. Sci. Statist. Comput. 9 (1988) 766–771.

[15] R. Freund, T. Huckle, Iterative solution of linear systems with low displacement rank by preconditioned conjugate gradient-type algorithms, in: XII Householder Symposium, Lake Arrowhead.

[16] I. Gohberg, T. Kailath, V. Olshevsky, Fast gaussian elimination with partial pivoting for matrices with displacement structure, Math. Comp. 64 (1995) 1557-1576.

[17] G. Heinig, A. Bojanczyk, Transformation techniques for Toeplitz and Toeplitz-plus-Hankel matrices, Linear Algebra Appl. 254 (1997) 193–226.

[18] G. Heinig, K. Rost, Representations of Toeplitz-plus-Hankel matrices using trigonometric transforms with application to fast matrix-vector multiplication, Linear Algebra Appl., in print.

[19] S. Holmgren, K. Otto, A framework for polynomial preconditioners based on fast transforms I: Theory, Preprint, Uppsala University, 1997.

[20] T. Huckle, Circulant/skewcirculant matrices as preconditioners for hermitian Toeplitz systems, in: IMACS Conference on Iterative methods in Linear Algebra, Brüssel, 1991.

[21] T. Huckle, Fast transforms for tridiagonal linear equations, Bit 34 (1994) 99–112.

[22] T. Huckle, A look-ahead algorithm for solving nonsymmetric linear Toeplitz equations, in: Proccedings of the Fifth SIAM conference, Snowbird, 1994, pp. 455–459.

[23] T. Huckle, Cauchy matrices and iterative methods for Toeplitz matrices, in: SPIE-95, 2563, San Diego, 1995, pp. 281–292.

[24] T. Kailath, V. Olshevsky, Displacement structure approach to discrete-trigonometric-transform based preconditioners of G. Strang type and of T. Chan type, Calcolo 33 (1996) 191–208.
[25] T. Kailath, V. Olshevsky, Diagonal pivoting for partially reconstructible Cauchy-like Matrices, with applications to Toeplitz-like linear equations and to boundary rational matrix interpolation problems, Linear Algebra Appl. 254 (1997) 251–302.
[26] T. Ku, C. Kuo, Spectral properties of preconditioned rational Toeplitz matrices, SIAM J. Matrix Anal. Appl. 14 (1993) 146–165.
[27] S.V. Parter, On the distribution of singular values of toeplitz matrices, Linear Algebra Appl. 80 (1986) 115–130.
[28] D. Potts, Schnelle Polynomtransformation und Vorkonditionierer für Toeplitz-Matrizen, Ph.D. Thesis, Universität Rostock, 1998, submitted.
[29] K. Rao, P. Yip, Discrete Cosine Transforms, Academic Press, Boston, 1990.
[30] Y. Saad, Iterative Methods for Sparse Linear Systems, Prindle, Weber & Schmidt, Boston, MA, 1996.
[31] G. Steidl, M. Tasche, A polynomial approach to fast algorithms for discrete Fourier-cosine and Fourier-sine transforms, Math. Comp. 56 (1991) 281–296.
[32] V. Strela, E.E. Tyrtyshnikov, Some generalizations of circulant preconditioner, Matrix Methods and Algorithms, IVM RAN Moscow, 1990, pp. 66–73.
[33] E.E. Tyrtyshnikov, Optimal and superoptimal circulant preconditioners, SIAM J. Matrix Anal. Appl. 13 (1992) 459–473.
[34] E.E. Tyrtyshnikov, A unifying approach to some old and new theorems on distribution and clustering, Linear Algebra Appl. 232 (1996) 1–43.
[35] E.E. Tyrtyshnikov, A. Yeremin, N. Zamarashkin, Clusters – preconditioners – convergence, Linear Algebra Appl. 263 (1997) 25–48.
[36] Z. Wang, Fast algorithms for the discrete W transform and for the discrete Fourier transform, IEEE Trans. Acoust. Speech Signal Process 32 (1984) 803–816.
[37] J. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965.