Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 418 (2006) 576–590
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Fast superoptimal preconditioning of multiindex Toeplitz matrices ☆

Cornelis van der Mee*, Giuseppe Rodriguez, Sebastiano Seatzu

Università di Cagliari, Dipartimento di Matematica e Informatica, viale Merello 92, 09123 Cagliari, Italy

Received 12 July 2005; accepted 27 February 2006
Available online 3 May 2006
Submitted by E.E. Tyrtyshnikov

Abstract

In this article three well-known methods of circulant preconditioning of finite multiindex Toeplitz linear systems, that is linear systems indexed by integers $i_1, \ldots, i_d$ with $0 \leqslant i_s < n_s$ $(s = 1, \ldots, d)$, are studied in detail. A general algorithm for the construction of the so-called superoptimal preconditioner is also given and it is shown that this procedure requires $\mathrm{O}(n_1 \cdots n_d \log(n_1 \cdots n_d))$ floating point operations.

AMS classification: 65F10; 15A69

Keywords: Circulant preconditioning; Multiindex matrices

1. Introduction

In the past two decades three circulant preconditioning methods have been introduced and implemented numerically to solve finite Toeplitz systems of the type

$$\sum_{j=0}^{n-1} a_{i-j} x_j = b_i, \quad i = 0, 1, \ldots, n-1,$$

☆ The research leading to this article was supported in part by MIUR under the COFIN grants no. 2004015437, and by INdAM-GNCS.
* Corresponding author. Tel.: +39 070 6755605; fax: +39 070 6755601.
E-mail addresses: cornelis@krein.unica.it (C. van der Mee), rodriguez@unica.it (G. Rodriguez), seatzu@unica.it (S. Seatzu).

0024-3795/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2006.02.034

where $A = (a_{i-j})_{i,j=0}^{n-1}$ is a complex Toeplitz matrix. These preconditioning methods consist of the determination of a nonsingular circulant matrix $C = (c_{i-j})_{i,j=0}^{n-1}[c_i = c_{i-n}, i = 1, \ldots, n-1]$ such that the ensuing system $C^{-1}A\mathbf{x} = C^{-1}\mathbf{b}$ $[\mathbf{x} = (x_i)_{i=0}^{n-1}, \mathbf{b} = (b_i)_{i=0}^{n-1}]$ can more easily be solved by iteration. They are commonly called Strang [9], optimal [5] and superoptimal [2,11,12] preconditioning, representing different strategies for finding the circulant preconditioner. Each preconditioning step of an $n \times n$ Toeplitz linear system by a circulant matrix can be implemented in $\mathrm{O}(n \log n)$ operations. A numerical investigation of the spectral properties of the Strang and the optimal preconditioning techniques in the one-level case is contained, e.g., in [10]. In [3] a survey of the results concerning the solution of Toeplitz linear systems by the preconditioned conjugate gradient method is given.

In this article we study circulant preconditioning of Toeplitz matrices A whose entries $a_{i,j}$ are indexed by $i, j \in E_n$, where

$$E_n = \{i = (i_1, \ldots, i_d) \in \mathbb{Z}^d : 0 \leqslant i_s < n_s \text{ for } s = 1, \ldots, d\} \tag{1}$$

and $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$. Here $\mathbb{Z}$ and $\mathbb{N}$ denote the sets of all integers and all positive integers, respectively. The matrix $A = (a_{i,j})_{i,j \in E_n}$ is called a *Toeplitz matrix* if $a_{i,j}$ only depends on $i - j = (i_1 - j_1, \ldots, i_d - j_d) \in \mathbb{Z}^d$.

When the d-index Toeplitz matrix $A = (a_{i-j})_{i,j \in E_n}$ is the tensor product of d one-index Toeplitz matrices, i.e. when

$$a_{i-j} = (A_1 \otimes \cdots \otimes A_d)_{ij} = \prod_{s=1}^{d} a_{i_s - j_s}^{(s)} \tag{2}$$

for $i = (i_1, \ldots, i_d) \in E_n$ and $j = (j_1, \ldots, j_d) \in E_n$, it is easily seen that circulant preconditioning can be implemented in

$$\mathrm{O}\left(\sum_{s=1}^{d} (n_1 \cdots n_{s-1} n_{s+1} \cdots n_d) n_s \log n_s\right) = \mathrm{O}(n_1 \cdots n_d \log(n_1 \cdots n_d))$$

operations, because the corresponding linear system is a *cascade* of d consecutive one-index Toeplitz systems.

Instead of d-level matrices, we employ d-index matrices $A = (a_{i,j})$ where $i = (i_1, \ldots, i_d)$ and $j = (j_1, \ldots, j_d)$ are d-tuples of integers (called multiindices) in E_n. In this way we do not need to introduce a lexicographical order on the multiindices i, j to keep track of the nesting pattern when using d-level matrices, which greatly simplifies the description when $d \geqslant 3$. Thus, instead of d-level Toeplitz matrices, which are $n_1 \times n_1$ block Toeplitz matrices whose entries are $(d-1)$-level Toeplitz matrices, which in turn are $n_2 \times n_2$ block Toeplitz matrices whose entries are $(d-2)$-level Toeplitz matrices, etc., we employ d-index Toeplitz matrices $A = (a_{i,j})$ whose elements $a_{i,j}$ only depend on $i - j \stackrel{\text{def}}{=} (i_1 - j_1, \ldots, i_d - j_d)$. By the same token, instead of d-level circulant matrices we use d-index circulant matrices $C = (c_{i,j})$ whose elements $c_{i,j}$ only depend on the unique integers $r_1, \ldots, r_d$ satisfying $0 \leqslant r_s < n_s$ for which $(i_s - j_s - r_s)/n_s$ is an integer $(s = 1, \ldots, d)$.

In [12] the author claims that the method proposed in the one-level case can easily be extended to the two-level case and requires $\mathrm{O}(n_1 n_2 \log(n_1 n_2))$ floating point operations. However, it is not clear how such an extension can be obtained except for tensor products of one-level matrices, because in the multilevel case the product of two lower/upper triangular Toeplitz matrices is not a Toeplitz matrix and this property is needed in the construction of the superoptimal one-level preconditioner. In [7], using a different technique, a method to construct the superoptimal preconditioner in the two-level case is given, but only if the matrix is hermitian. As a result, in the literature there is no general method to construct the superoptimal preconditioner in the d-level case in $\mathrm{O}(n\log(n))$ floating point operations, where $n = n_1 \cdots n_d$. Motivated by this gap in the literature and by the apparent interest in using superoptimal preconditioners ([8], also [7]), we have developed a method to compute the superoptimal preconditioner in $\mathrm{O}(n_1 \cdots n_d \log(n_1 \cdots n_d))$ operations.

Putting E_n into the obvious $(1, 1)$-correspondence with the finite additive group

$$G = \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_d}, \tag{3}$$

where $\mathbb{Z}_{n_s}$ is the additive group of integers modulo n_s, we call $C = (c_{i,j})_{i,j\in E_n}$ a *circulant matrix* if its entries $c_{i,j}$ only depend on the difference $i - j$ when taken in the group G, which makes it into a special kind of Toeplitz matrix. We shall consider the above three types of circulant preconditioning of Toeplitz systems and prove that they can be implemented in $\mathrm{O}(n_1 \cdots n_d \log(n_1 \cdots n_d))$ operations, also for multiindex Toeplitz matrices which do not have the tensor product structure.

It is straightforward to generalize the Strang and optimal preconditioning methods from the one-index to the d-index case (see Section 3) and to estimate the number of operations required for their numerical implementation. The superoptimal preconditioner can also be introduced in the d-index case in a straightforward way, but its numerical implementation in $\mathrm{O}(n_1 \cdots n_d \log(n_1 \cdots n_d))$ is more involved. Our method represents a generalization to the d-index case of the method by Chan et al. [2] and Tismenetsky [11], where a Toeplitz matrix is written as the sum of a circulant and a skew-circulant matrix. We show that our method requires $\mathrm{O}(n_1 \cdots n_d \log(n_1 \cdots n_d))$ operations, where the order constant is asymptotically $\mathrm{O}(4^d)$ as $d \to \infty$.

Multiindex circulant preconditioning is a useful tool in many structured or *almost* structured problems of large dimension, such as the discretization of integral equations with convolutive kernels or the solution of partial differential equations on regular domains by finite differences. Such problems occur in many applicative situations, e.g. in various kinds of medical tomography (CT, PET, SPECT) modelled by the Radon trasform, whose discretization in particular physical and geometrical situations leads to multiindex structured matrices, and in many engineering problems, like geophysical prospection and remote sensing, which need to be solved in three dimensions.

2. Basic formalism

2.1. The one-index case

Let us compile some well-known facts [6].

The $n \times n$ circulant matrices $C = (c_{i-j})_{i,j=0}^{n-1} [c_i = c_{i-n}, i = 1, \ldots, n-1]$ form a C^*-algebra with unit element with respect to the usual matrix addition, multiplication and conjugate transposition. Fixing a primitive nth root of unity η (i.e., $\eta^n = 1$ but $\eta^i \neq 1$ for $i = 1, \ldots, n-1$), the Fourier matrix

$$\mathscr{F}_\eta = \frac{1}{\sqrt{n}} (\eta^{ij})_{i,j=0}^{n-1}$$

is unitary and diagonalizes C:

$$C\mathscr{F}_\eta = \mathscr{F}_\eta \mathrm{diag}(\hat{C}(\eta^{-j}))_{j=0}^{n-1}, \tag{4}$$

where

$$\hat{C}(\zeta) \stackrel{\text{def}}{=} \sum_{i=0}^{n-1} \zeta^i c_i \tag{5}$$

is the *generating polynomial* of C. The eigenvalues of C are exactly the values of this polynomial at the nth roots of unity. The Fourier matrix $\mathscr{F}_\eta$ defines a norm preserving isometry between the C^*-algebra of circulant matrices and $\mathbb{C}^n$ (with maximum norm $\|\zeta\| = \max(|\zeta_1|, \ldots, |\zeta_n|)$, $\zeta = (\zeta_1, \ldots, \zeta_n)$).

The $n \times n$ skew-circulant matrices $S = (s_{i-j})_{i,j=0}^{n-1}$ $[s_i = -s_{i-n}, i = 1, \ldots, n-1]$ also form a C^*-algebra with unit element. Using the above η and fixing a root ω of the equation $\omega^n = -1$, we arrive at the following diagonalization of S:

$$SD_\omega\mathscr{F}_\eta = D_\omega\mathscr{F}_\eta \operatorname{diag}(\hat{S}(\omega^{-1}\eta^{-j}))_{j=0}^{n-1}, \tag{6}$$

where $\hat{S}$ is given in terms of $(s_i)_{i=0}^{n-1}$ by (5) and $D_\omega = \operatorname{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1})$ is a unitary diagonal matrix. The eigenvalues of S are exactly the values of the generating polynomial at the roots of $z^n = -1$. The matrix $D_\omega\mathscr{F}_\eta$ defines a norm preserving isometry between the C^*-algebra of skew-circulant matrices and $\mathbb{C}^n$ endowed with the maximum norm.

Eqs. (4) and (6) yield the obvious inversion formulas for circulant and skew-circulant matrices

$$C^{-1} = \begin{cases} \mathscr{F}_\eta \operatorname{diag}\left(\frac{1}{\hat{C}(\eta^{-j})}\right)_{j=0}^{n-1} \mathscr{F}_\eta^*, & \text{circulant case,} \\ D_\omega\mathscr{F}_\eta \operatorname{diag}\left(\frac{1}{\hat{C}(\omega^{-1}\eta_1^{-j})}\right)_{j=0}^{n-1} \mathscr{F}_\eta^* D_\omega^*, & \text{skew-circulant case,} \end{cases}$$

provided, of course, that $\hat{C}(\zeta) \neq 0$ for all roots of $\zeta^n - 1 = 0$ (circulant case) or $\zeta^n + 1 = 0$ (skew-circulant case). In algebraic terms, the inverse of a circulant (skew-circulant) matrix $C = (c_{i-j})_{i,j=0}^{n-1}$ is the circulant (skew-circulant) matrix $D = (d_{i-j})_{i,j=0}^{n-1}$ for which $\hat{C}(z)\hat{D}(z) - 1$ is divisible by $z^n - 1$ ($z^n + 1$).

Every Toeplitz matrix $A = (a_{i-j})_{i,j=0}^{n-1}$ can be written as the sum $A = C + S$ of a circulant matrix $C = (c_{i-j})_{i,j=0}^{n-1}$ and a skew-circulant matrix $S = (s_{i-j})_{i,j=0}^{n-1}$ as follows:

$$c_0 = a_0, \quad s_0 = 0, \quad c_i = c_{i-n} = \frac{a_i + a_{i-n}}{2}, \quad s_i = -s_{i-n} = \frac{a_i - a_{i-n}}{2}, \tag{7}$$

where $i = 1, \ldots, n-1$. The decomposition $a_0 = c_0 + s_0$ is somewhat arbitrary, but here we decide to choose $c_0 = a_0$ and $s_0 = 0$. As a result, C is selfadjoint and S is skew-selfadjoint whenever A is selfadjoint.

2.2. The d-index case

Let $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$. Then $\zeta = (\zeta_1, \ldots, \zeta_d) \in \mathbb{C}^d$ is called an nth *root of unity* if $\zeta_1^{n_1} = \cdots = \zeta_d^{n_d} = 1$, and a *primitive* nth *root of unity* if

$$\{\zeta_s^{m_s} : 1 \leqslant m_s < n_s \text{ and } s = 1, \ldots, d\} \cap \{1\} = \emptyset.$$

For any complex vector ζ we write $\zeta^p = \zeta_1^{p_1} \cdots \zeta_d^{p_d}$ if $p = (p_1, \ldots, p_d) \in \mathbb{Z}^d$. Then it is easily seen that the additive group G defined by (3) is isomorphic to the multiplicative group of nth roots of unity, where an isomorphism is given by the map

$$i = (i_1, \ldots, i_d) \mapsto (\eta_1^{i_1}, \ldots, \eta_d^{i_d})$$

for any primitive nth root of unity η.

Given $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$, by a Toeplitz matrix of multiorder n we mean a matrix $A = (a_{i,j})_{i,j \in E_n}$ indexed by the set E_n in (1) whose entries only depend on $i - j$: $a_{i,j} = a_{i-j}$. By a circulant matrix of multiorder n we mean a Toeplitz matrix of multiorder n, $C = (c_{i-j})_{i,j \in E_n}$, for which the following d relations hold:

$$c_{(i_1,\ldots,i_{s-1},i_s-n_s,i_{s+1},\ldots,i_d)} = c_{(i_1,\ldots,i_d)}, \quad i = (i_1, \ldots, i_d) \in E_n, \ \ i_s > 0, \tag{8}$$

where $s = 1, \ldots, d$. Using the natural $(1,1)$-correspondence between the elements of E_n and those of the group G in (3), we see that the circulant matrices of multiorder n are exactly those matrices $C = (c_{i,j})_{i,j \in G}$ indexed by G for which the elements only depend on the difference $i - j$ in G. Thus we may alternatively write $C = (c_{i-j})_{i,j \in G}$.

The circulant matrices $C = (c_{i-j})_{i,j \in E_n}$ form a C^*-algebra with unit element with respect to the usual matrix addition, multiplication and conjugate transposition. Fixing a primitive nth root of unity $\eta = (\eta_1, \ldots, \eta_d)$, the Fourier matrix[1]

$$\mathscr{F}_\eta = \frac{1}{\sqrt{n_1 \cdots n_d}} (\eta^{i \circ j})_{i,j \in E_n} = \left(\prod_{s=1}^{d} \frac{1}{\sqrt{n_s}} \eta_s^{i_s j_s} \right)_{i,j \in E_n}$$

is unitary and diagonalizes C:

$$C \mathscr{F}_\eta = \mathscr{F}_\eta (\hat{C}(\eta_1^{-j_1}, \ldots, \eta_d^{-j_d}))_{j \in E_n},$$

where

$$\hat{C}(\zeta) \stackrel{\text{def}}{=} \sum_{i \in E_n} \zeta^i c_i = \sum_{i_1=0}^{n_1-1} \cdots \sum_{i_d=0}^{n_d-1} \zeta_1^{i_1} \cdots \zeta_d^{i_d} c_{(i_1,\ldots,i_d)} \tag{9}$$

is the d-variate *generating polynomial* of C. The eigenvalues of C are exactly the values of this polynomial at the nth roots of unity. The Fourier matrix $\mathscr{F}_\eta$ defines a norm preserving isometry between the C^*-algebra of circulant matrices of multiorder n and $\mathbb{C}^{n_1 \cdots n_d}$ endowed with the maximum norm.

Given $\sigma \in \{0, 1\}^d$, by a *σ-circulant matrix* of multiorder n we mean a Toeplitz matrix $C = (c_{i-j})_{i,j \in E_n}$ of multiorder n which satisfies the condition

$$c_k = 0 \quad \text{if } \exists\, s : k_s = 0 \text{ and } \sigma_s = 1$$

and the symmetry relation

$$c_{k-(\tau \circ n)} = (-1)^{\tau \cdot \sigma} c_k$$

for all $k \in E_n$ and $\tau \in \{0, 1\}^d$ with $\tau_s = 0$ whenever $k_s = 0$. Clearly, the σ-circulant matrices with $\sigma = (0, \ldots, 0)$ are exactly the circulant matrices of multiindex n.

Example 1. Let $d = 2$, $n = (n_1, n_2) \in \mathbb{N}^2$, $0 < k_1 < n_1$ and $0 < k_2 < n_2$. Then for each $\sigma \in \{0, 1\}^2$ the entries of the first column in the following table are equal to the corresponding entries in the column labelled by σ

[1] Throughout we let $z \circ w = (z_1 w_1, \ldots, z_d w_d)$ stand for the Schur product of the vectors $z = (z_1, \ldots, z_d)$ and $w = (w_1, \ldots, w_d)$ in $\mathbb{C}^d$, and $z \cdot w = z_1 w_1 + \cdots + z_d w_d$ for their inner product.

	$\sigma=(0,0)$	$\sigma=(1,0)$	$\sigma=(0,1)$	$\sigma=(1,1)$
$c^{(\sigma)}_{(k_1-n_1,k_2)}$	$c_{(k_1,k_2)}$	$-c_{(k_1,k_2)}$	$c_{(k_1,k_2)}$	$-c_{(k_1,k_2)}$
$c^{(\sigma)}_{(k_1,k_2-n_2)}$	$c_{(k_1,k_2)}$	$c_{(k_1,k_2)}$	$-c_{(k_1,k_2)}$	$-c_{(k_1,k_2)}$
$c^{(\sigma)}_{(k_1-n_1,k_2-n_2)}$	$c_{(k_1,k_2)}$	$-c_{(k_1,k_2)}$	$-c_{(k_1,k_2)}$	$c_{(k_1,k_2)}$
$c^{(\sigma)}_{(k_1-n_1,0)}$	$c_{(k_1,0)}$	$-c_{(k_1,0)}$	0	0
$c^{(\sigma)}_{(0,k_2-n_2)}$	$c_{(0,k_2)}$	0	$-c_{(0,k_2)}$	0
$c^{(\sigma)}_{(0,0)}$	$c_{(0,0)}$	0	0	0

Now let $\omega=(\omega_1,\ldots,\omega_d)$ be a vector of fixed entries $\omega_1,\ldots,\omega_d$ such that $\omega_1^{n_1}=\cdots=\omega_d^{n_d}=-1$, and let $\eta=(\eta_1,\ldots,\eta_d)$ be a primitive nth root of unity. Given $\sigma\in\{0,1\}^d$, we define the diagonal matrix

$$D^{(\sigma)}_\omega=\operatorname{diag}(\omega_1^{\sigma_1 j_1}\cdots\omega_d^{\sigma_d j_d})_{j\in E_n}.$$

Then a σ-circulant matrix C can be diagonalized as follows:

$$CD^{(\sigma)}_\omega\mathscr{F}_\eta=D^{(\sigma)}_\omega\mathscr{F}_\eta\operatorname{diag}(\hat{C}(\omega_1^{-\sigma_1}\eta_1^{-j_1},\ldots,\omega_d^{-\sigma_d}\eta_d^{-j_d}))_{j\in E_n},\tag{10}$$

where $\hat{C}$ is given by (9) and $D^{(\sigma)}_\omega\mathscr{F}_\eta$ is a unitary matrix. Thus the eigenvalues of a σ-circulant matrix of multiorder n are exactly the values of the generating polynomial at the zeros of the d-variate polynomial $\mathscr{P}^{(\sigma)}_n(z)\stackrel{\text{def}}{=}\prod_{s=1}^d(z^{n_s}-(-1)^{\sigma_s})$. It is now easily seen that the σ-circulant matrices of multiorder n form a C^*-algebra with unit element, where $D^{(\sigma)}_\omega\mathscr{F}_\eta$ defines a norm preserving isomorphism between the C^*-algebra of σ-circulant matrices and $\mathbb{C}^{n_1\cdots n_d}=\mathbb{C}^{n_1}\times\cdots\times\mathbb{C}^{n_d}$ endowed with the maximum norm.

Eq. (10) yields the obvious inversion formula for σ-circulant matrices

$$C^{-1}=D^{(\sigma)}_\omega\mathscr{F}_\eta\operatorname{diag}\left(\frac{1}{\hat{C}(\omega_1^{-\sigma_1}\eta_1^{-j_1},\ldots,\omega_d^{-\sigma_d}\eta_d^{-j_d})}\right)_{j\in E_n}\mathscr{F}_\eta^*D^{(\sigma)*}_\omega,$$

provided, of course, that $\hat{C}(\zeta)\neq 0$ for all roots of $\mathscr{P}^{(\sigma)}_n(z)=0$. In algebraic terms, the inverse of a σ-circulant matrix $C=(c_{i-j})_{i,j\in E_n}$ of multiorder n is the σ-circulant matrix $D=(d_{i-j})_{i,j\in E_n}$ for which $\hat{C}(z)\hat{D}(z)-1$ is divisible by the d-variate polynomial $\mathscr{P}^{(\sigma)}_n(z)$.

Let us now generalize a result regarding one-index Toeplitz matrices [11]. To this aim it is useful to introduce the set

$$\begin{aligned}F_n=E_n-E_n=\big\{&(i_1,\ldots,i_d)\in\mathbb{Z}^d:-(n_1-1)\leqslant i_1\leqslant(n_1-1),\ldots,\\&-(n_d-1)\leqslant i_d\leqslant(n_d-1)\big\}.\end{aligned}\tag{11}$$

Theorem 2. *Every Toeplitz matrix $A=(a_{i-j})_{i,j\in E_n}$ can be written in the form*

$$A=\sum_{\sigma\in\{0,1\}^d}C^{(\sigma)},\tag{12}$$

where $C^{(\sigma)}$ is a σ-circulant matrix. More precisely,

$$c_k^{(\sigma)} = \begin{cases} \left(\frac{1}{2}\right)^{\#\{s:k_s\neq 0\}} \displaystyle\sum_{\substack{\tau\in\{0,1\}^d\\ k_s=0\Rightarrow\tau_s=0}} (-1)^{\tau\cdot\sigma} a_{k-(\tau\circ n)}, & \sigma_s = 0 \textit{ whenever } k_s = 0,\\ 0, & \exists s : k_s = 0 \textit{ and } \sigma_s = 1. \end{cases} \tag{13}$$

In particular,

$$c_{(0,\ldots,0)}^{(\sigma)} = \begin{cases} a_{(0,\ldots,0)}, & \sigma = (0,\ldots,0),\\ 0, & \sigma \neq (0,\ldots,0). \end{cases}$$

Proof. Let $\sigma = (\boldsymbol{\sigma}, \sigma_d)$, $n = (\boldsymbol{n}, n_d)$, $\tau = (\boldsymbol{\tau}, \tau_d)$, $k = (\boldsymbol{k}, k_d)$, $E_n = E_{\boldsymbol{n}} \times E_{n_d}$, and $F_n = E_n - E_n = F_{\boldsymbol{n}} \times F_{n_d}$, and let

$$A = (a_{i-j})_{i,j\in E_n} = (a_{(\boldsymbol{i}-\boldsymbol{j},i_d-j_d)})_{\boldsymbol{i},\boldsymbol{j}\in E_{\boldsymbol{n}};i_d,j_d\in E_{n_d}}$$

be a d-index Toeplitz matrix. Then, assuming Theorem 2 to be true for $(d-1)$-index Toeplitz matrices, we have for every $k_d \in F_{n_d} = E_{n_d} - E_{n_d}$

$$(a_{(\boldsymbol{i}-\boldsymbol{j},k_d)})_{\boldsymbol{i},\boldsymbol{j}\in E_{\boldsymbol{n}}} = \sum_{\boldsymbol{\sigma}\in\{0,1\}^{d-1}} C_{[k_d]}^{(\boldsymbol{\sigma})}, \tag{14}$$

where $C_{[k_d]}^{(\boldsymbol{\sigma})} = (c_{(\boldsymbol{i}-\boldsymbol{j},[k_d])}^{(\boldsymbol{\sigma})})_{\boldsymbol{i},\boldsymbol{j}\in E_{\boldsymbol{n}}}$ is a $\boldsymbol{\sigma}$-circulant matrix.

Thus for any $k_d \in F_{n_d}$ we have

$$c_{(\boldsymbol{k},[k_d])}^{(\boldsymbol{\sigma})} = \begin{cases} \left(\frac{1}{2}\right)^{\#\{s<d:k_s\neq 0\}} \displaystyle\sum_{\substack{\boldsymbol{\tau}\in\{0,1\}^{d-1}\\ k_s=0\Rightarrow\tau_s=0}} (-1)^{\boldsymbol{\tau}\cdot\boldsymbol{\sigma}} a_{(\boldsymbol{k}-\boldsymbol{\tau}\circ\boldsymbol{n},k_d)}, & \sigma_s = 0 \text{ whenever } k_s = 0, \text{ and } s < d,\\ 0, & \exists s \in \{1,\ldots,d-1\} : k_s = 0 \text{ and } \sigma_s = 1. \end{cases}$$

We now define

$$c_k^{(\sigma)} = c_{(\boldsymbol{k},k_d)}^{(\boldsymbol{\sigma},\sigma_d)} = \begin{cases} c_{(\boldsymbol{k},[0])}^{(\boldsymbol{\sigma})}, & k_d = 0 \text{ and } \sigma_d = 0,\\ 0, & k_d = 0 \text{ and } \sigma_d = 1,\\ \dfrac{c_{(\boldsymbol{k},[k_d])}^{(\boldsymbol{\sigma})} + (-1)^{\sigma_d} c_{(\boldsymbol{k},[k_d-n_d])}}{2}, & k_d \neq 0 \text{ and } \sigma_d \in \{0,1\}. \end{cases} \tag{15}$$

Eqs. (7), (14) and (15) are now easily seen to imply (12) and (13). $\square$

3. Circulant preconditioning

In this section we generalize the three traditional methods of circulant preconditioning of Toeplitz systems from the one-index to the d-index case.

Strang-type preconditioning has been developed to solve symmetric Toeplitz systems by the conjugate gradient method with circulant preconditioning, where the circulant preconditioner has the same elements on a suitable band as the given Toeplitz matrix. In order to generalize Strang-type preconditioning to (non-necessarily symmetric) Toeplitz systems of multiorder $n = (n_1,\ldots,n_d) \in \mathbb{N}^d$, we call a subset E of F_n, with F_n defined by (11), an *admissible band* if $i, j \in E$ and $((i_s - j_s)/n_s) \in \mathbb{Z}$ $(s = 1,\ldots,d)$ imply $i = j$. In that case we can associate to any Toeplitz matrix $A = (a_{i-j})_{i,j\in E_n}$ a circulant matrix $S = (s_{i-j})_{i,j\in E_n}$ satisfying $s_i = a_i$ for each $i \in E$ by defining

$$s_i = \begin{cases} a_j, & i \in E \text{ and } ((i_s - j_s)/n_s) \in \mathbb{Z}, \\ 0, & \text{otherwise.} \end{cases}$$

Here we note that s_i is well-defined if E is an admissible band. We also note that the matrix S is real whenever A is real, real symmetric whenever A is real symmetric, and hermitian whenever A is hermitian.

In the one-index case the admissible bands are exactly the subsets E of F_n such that no two distinct elements $i, j \in E$ satisfy $((i - j)/n) \in \mathbb{Z}$. Among the admissible bands are the sets $E = \{-m, -m + 1, \ldots, m - 1, m\}$ for which $2m \leqslant n - 1$. Strang preconditioning is usually associated with the admissible band $\{m - n + 1, m - n + 2, \ldots, m - 1, m\}$ where $m = \lfloor n/2 \rfloor$ and $\lfloor x \rfloor$ denotes the largest integer m such that $m \leqslant x$.

In the d-index case the admissible bands are precisely the subsets E of F_n such that no distinct elements $i, j \in E$ satisfy $((i_s - j_s)/n_s) \in \mathbb{Z}$. Among the admissible bands are all of the sets $E = \{i \in \mathbb{Z}^d : -m_s \leqslant i_s \leqslant m_s \ (s = 1, \ldots, d)\}$ where $2m_s \leqslant n_s - 1$ $(s = 1, \ldots, d)$. The natural generalization of Strang preconditioning is that associated with the admissible band $\{i \in \mathbb{Z}^d : m_s - n_s + 1 \leqslant i_s \leqslant m_s \ (s = 1, \ldots, d)\}$ where $m_s = \lfloor n_s/2 \rfloor (s = 1, \ldots, d)$.

The construction of S requires no computation. This preconditioner was first introduced in [9] and its properties in the one-index case were investigated in [4].

In order to study optimal preconditioning of d-index matrices, we index such matrices by $i, j \in G$, where G is given by (3), and reformulate here for a complex matrix a result already proved in the real case in [12].

Theorem 3. *Let $A = (a_{i,j})_{i,j \in G}$ be a complex G-indexed matrix. Then the complex circulant matrix $C = (c_{i-j})_{i,j \in G}$ closest to A in the Frobenius norm is given by*

$$c_p = \frac{1}{\#G} \sum_{\substack{(i,j) \in G \times G \\ i-j=p \text{ in } G}} a_{i,j}, \quad p \in G. \tag{16}$$

The matrix C is real whenever A is real, real symmetric whenever A is real symmetric, hermitian whenever A is hermitian, and positive definite if A is positive definite.

Proof. The proof is elementary and requires minimizing a quadratic polynomial of $2(\#G)$ real variables. $\square$

The so-called optimal preconditioner (or Chan's preconditioner) is the circulant matrix C defined by (16), which we will henceforth denote by $\mathscr{C}_A$. When invertible, it allows one to replace the linear system $A\mathbf{x} = \mathbf{b}$ for G-indexed column vectors $\mathbf{x}$ and $\mathbf{b}$ by the linear system $\mathscr{C}_A^{-1} A\mathbf{x} = \mathscr{C}_A^{-1}\mathbf{b}$ which is supposedly better conditioned than $A\mathbf{x} = \mathbf{b}$. It appeared for the first time in [5]; its spectral properties were studied in [1] and in many subsequent papers.

We now give a characterization of the superoptimal preconditioner. Let $n = (n_1, \ldots, n_d)$, define G as in (3), and put $\#G = n_1 \cdots n_d$.

Theorem 4. *Suppose $A = (a_{i,j})_{i,j \in G}$ is a complex matrix such that for each nth root of unity ζ there exists at least one $j \in G$ such that $\sum_{p \in G} a_{p,j} \zeta^{-p} \neq 0$. Then the complex circulant matrix $D = (d_{i-j})_{i,j \in G}$ such that DA is closest to the identity matrix in the Frobenius norm is unique and is given by the solution of the linear system*

$$\sum_{k \in G} T_{p,k} d_k = g_p, \quad p \in G, \tag{17}$$

where

$$T_{p,k} = \frac{1}{\#G} \sum_{(i,j)\in G\times G} \overline{a_{i-p,j}} a_{i-k,j} = [\mathscr{C}_{AA^*}]_{p-k}, \tag{18}$$

$$g_p = \frac{1}{\#G} \sum_{i\in G} \overline{a_{i-p,i}} = [\mathscr{C}_{A^*}]_p. \tag{19}$$

The matrix D is real whenever A is real, real symmetric whenever A is real symmetric, hermitian whenever A is hermitian and positive definite if A is positive definite. Further, D is nonsingular if and only if $\mathscr{C}_{A^*}$ is nonsingular.

Proof. The proof of (17) follows from minimizing $\|I - DA\|_F^2$ with respect to the first column $(d_p)_{p\in G}$ of D.

Using Theorem 3 it is then immediate that $T = \mathscr{C}_{AA^*}$, which turns T into a circulant matrix. Also the right-hand sides g_p of (17) are the elements in the first column of $\mathscr{C}_{A^*}$. The identity

$$\frac{1}{\#G} \sum_{(p,k)\in G\times G} T_{p,k}\xi_p\overline{\xi_k} = \sum_{(i,j)\in G\times G} \left| \frac{1}{\#G} \sum_{p\in G} \overline{a_{i-p,j}}\xi_p \right|^2,$$

where ξ_p are complex numbers indexed by $p \in G$, implies the positive definiteness of T. Moreover, the generating polynomial of T is given by

$$\hat{T}(\zeta) = \sum_{(i,j)\in G\times G} \left| \frac{1}{\#G} \sum_{p\in G} \overline{a_{i-p,j}}\zeta^{p+j-i} \right|^2 = \frac{1}{\#G} \sum_{j\in G} \left| \sum_{p\in G} \overline{a_{p,j}}\zeta^{j-p} \right|^2.$$

Consequently, the linear system (17) is uniquely solvable, unless $\hat{T}(\zeta) = 0$ for some nth root of unity ζ, and this can only happen if for some nth root of unity ζ the discrete Fourier transforms of all columns of A vanish.

If $\hat{T}(\zeta) = 0$ for some nth root of unity, then it may happen that (17) does not have any solution. This occur, e.g., if A is a singular circulant matrix. In this case (17) reduces to finding a circulant matrix D such that $AA^*D = A^*AD = A^*$, which is impossible unless A is invertible. □

4. Computational complexity

In this section we prove that superoptimal circulant preconditioning of a Toeplitz matrix of multiorder $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$ can be implemented in $\mathrm{O}(n_1 \cdots n_d \log(n_1 \cdots n_d))$ operations.

Let $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$ and let G be given by (3). Given a matrix $A = (a_{i,j})_{i,j\in G}$ indexed by G, its optimal (resp., superoptimal) circulant preconditioner is the circulant matrix $C = (c_{i-j})_{i,j\in G}$ (resp., $D = (d_{i-j})_{i,j\in G}$) of multiorder n which minimizes $\|C - A\|_F$ (resp., $\|I - DA\|_F$). Using that $\mathscr{F}_\eta^* C\mathscr{F}_\eta$ and $\mathscr{F}_\eta^* D\mathscr{F}_\eta$ are diagonal matrices for a given primitive nth root of unity η, we can define $B = \mathscr{F}_\eta^* A\mathscr{F}_\eta$ and reformulate the optimal (resp., superoptimal) circulant preconditioning problem as follows: Find a diagonal matrix $\Gamma = \mathrm{diag}(\gamma_i)_{i\in G}$ which minimizes $\|\Gamma - B\|_F$ (resp., $\|I - \Gamma B\|_F$) and put $C = \mathscr{F}_\eta \Gamma \mathscr{F}_\eta^*$ (resp., $D = \mathscr{F}_\eta \Gamma \mathscr{F}_\eta^*$). Minimizing $\|\Gamma - B\|_F^2$ (resp., $\|I - \Gamma B\|_F^2$) as a function of the real and imaginary parts of $\gamma_i (i \in G)$, we obtain

$$\gamma_i = b_{ii}, \quad i \in G,$$

in the case of the optimal preconditioner, with the minimal Frobenius distance given by $\|B - \mathrm{diag}(b_{ii})_{i\in G}\|_F$, and

$$\gamma_i = \frac{\overline{b_{ii}}}{\sum_{j\in G}|b_{ij}|^2} = \frac{\overline{b_{ii}}}{[BB^*]_{ii}}, \quad i \in G, \tag{20}$$

in the case of the superoptimal preconditioner. We analyze in detail the computation of the latter expression.

4.1. The one-index case

In the one-index case (where $G = \mathbb{Z}_n$ for some $n \in \mathbb{N}$) we apply (4), (6) and (7) to prove that

$$B = \mathscr{F}_\eta^* A \mathscr{F}_\eta = \Delta_{\hat{C}} + \mathscr{E}_\omega \Delta_{\hat{S}} \mathscr{E}_\omega^*, \tag{21}$$

where

$$\begin{aligned} \Delta_{\hat{C}} &= \mathrm{diag}(\hat{C}(\eta^{-j}))_{j=0}^{n-1}, \\ \Delta_{\hat{S}} &= \mathrm{diag}(\hat{S}(\omega^{-1}\eta^{-j}))_{j=0}^{n-1}, \\ \mathscr{E}_\omega &= (e_{i-j})_{i,j=0}^{n-1} = \mathscr{F}_\eta^* D_\omega \mathscr{F}_\eta, \end{aligned} \tag{22}$$

and

$$e_{i-j} = \frac{1}{n}\sum_{k=0}^{n-1}(\omega\eta^{j-i})^k = \frac{2}{n(1-\omega\eta^{j-i})},$$

which implies

$$b_{ii} = \hat{C}(\eta^{-i}) + \sum_{j=0}^{n-1}|e_{i-j}|^2 \hat{S}(\omega^{-1}\eta^{-j}), \tag{23}$$

$$[BB^*]_{ii} = |\hat{C}(\eta^{-i})|^2 + \sum_{j=0}^{n-1}|e_{i-j}|^2\Big[|\hat{S}(\omega^{-1}\eta^{-j})|^2 + 2\mathrm{Re}\Big\{\hat{C}(\eta^{-i})\overline{\hat{S}(\omega^{-1}\eta^{-j})}\Big\}\Big]. \tag{24}$$

Substituting (23) and (24) into (20) and putting $D = \mathscr{F}_\eta \Gamma \mathscr{F}_\eta^*$ we obtain an expression for the superoptimal circulant preconditioner that can be evaluated in $\mathrm{O}(n\log n)$ operations, as shown by different means by Chan et al. [2] and Tismenetsky [11].

4.2. The d-index case

Put $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$. In the d-index case (where G is given by (3)) we apply (12) to derive the expression

$$B = \mathscr{F}_\eta^* A \mathscr{F}_\eta = \sum_{\sigma\in\{0,1\}^d} \mathscr{E}_\omega^{(\sigma)} \Delta^{(\sigma)} \mathscr{E}_\omega^{(\sigma)*}, \tag{25}$$

where

$$\Delta^{(\sigma)} = \mathrm{diag}(\hat{C}^{(\sigma)}(\omega_1^{-\sigma_1}\eta_1^{-j_1}, \ldots, \omega_d^{-\sigma_d}\eta_d^{-j_d}))_{j\in E_n}, \tag{26}$$

and the matrix

$$\mathscr{E}_\omega^{(\sigma)} = (e_{i-j}^{(\sigma)})_{i,j\in G} = \mathscr{F}_\eta^* D_\omega^{(\sigma)} \mathscr{F}_\eta$$

is unitary circulant with entries given by

$$e_{(i_1-j_1,\ldots,i_d-j_d)}^{(\sigma)} = \prod_{s=1}^{d} (\mathscr{E}_{\omega_s}^{(\sigma_s)})_{i_s j_s} = \prod_{s:\sigma_s=1} \frac{2}{n_s(1-\omega_s\eta_s^{j_s-i_s})},$$

where $\mathscr{E}_{\omega_s}^{(0)} = I_{n_s}$ and $\mathscr{E}_{\omega_s}^{(1)} = \mathscr{E}_{\omega_s}$ (see (22)).

As a result of (25) we have

$$b_{ii} = \sum_{\sigma\in\{0,1\}^d} \sum_{j\in G} |e_{i-j}^{(\sigma)}|^2 \Delta_j^{(\sigma)}, \tag{27}$$

where $\Delta_j^{(\sigma)} = \hat{C}^{(\sigma)}(\omega_1^{-\sigma_1}\eta_1^{-j_1},\ldots,\omega_d^{-\sigma_d}\eta_d^{-j_d})$ and

$$\begin{aligned}[BB^*]_{ii} &= \sum_{\sigma,\tau\in\{0,1\}^d} \sum_{j,k\in G} e_{i-j}^{(\sigma)} \Delta_j^{(\sigma)} \big[\mathscr{E}_\omega^{(\sigma)*} \mathscr{E}_\omega^{(\tau)}\big]_{j-k} \overline{\Delta_k^{(\tau)}}\, \overline{e_{i-k}^{(\tau)}} \\ &= \sum_{\sigma,\tau\in\{0,1\}^d} \sum_{k\in G} \big[\mathscr{E}_\omega^{(\sigma)} \Delta^{(\sigma)} \mathscr{E}_\omega^{(\neg\sigma)}\big]_{ik} \overline{[\mathscr{E}_\omega^{(\tau)} \Delta^{(\tau)} \mathscr{E}_\omega^{(\neg\tau)}]_{ik}}. \end{aligned} \tag{28}$$

Here $\neg\sigma = (1-\sigma_1,\ldots,1-\sigma_d)$ for all $\sigma \in \{0,1\}^d$. The second line of (28) follows from the first line by observing that

$$\begin{aligned}\big[\mathscr{E}_\omega^{(\sigma)*}\mathscr{E}_\omega^{(\tau)}\big]_{j-k} &= \bigotimes_{\substack{s=1,\ldots,d\\ \sigma_s\neq\tau_s}} \big[\mathscr{E}_{\omega_s}^{(\tau_s)}\mathscr{E}_{\omega_s}^{(\sigma_s)*}\big]_{j_s-k_s} = \bigotimes_{\substack{s=1,\ldots,d\\ \sigma_s\neq\tau_s}} \big[\mathscr{E}_{\omega_s}^{(1-\sigma_s)}\mathscr{E}_{\omega_s}^{(1-\tau_s)*}\big]_{j_s-k_s} \\ &= \bigotimes_{s=1}^{d} \big[\mathscr{E}_{\omega_s}^{(1-\sigma_s)}\mathscr{E}_{\omega_s}^{(1-\tau_s)*}\big]_{j_s-k_s} = \big[\mathscr{E}_\omega^{(\neg\sigma)}\mathscr{E}_\omega^{(\neg\tau)*}\big]_{j-k}.\end{aligned}$$

Substituting (27) and (28) into (20) and putting $D = \mathscr{F}_\eta \Gamma \mathscr{F}_\eta^*$ we obtain an expression for the superoptimal circulant preconditioner that can be evaluated in $\mathrm{O}(n_1\cdots n_d \log(n_1\cdots n_d))$ operations, as we will now show in more detail.

Indeed, putting $\overline{\sigma} = \{s\in\{1,\ldots,d\} : \sigma_s = 1\}$ and $\underline{\sigma} = \{s\in\{1,\ldots,d\} : \sigma_s = 0\}$ we have

$$\big[\mathscr{E}_\omega^{(\sigma)}\Delta^{(\sigma)}\mathscr{E}_\omega^{(\neg\sigma)}\big]_{ik} = e_{i-k}^{(1,\ldots,1)} \Delta_{\mathrm{mix}_\sigma(i,k)}^{(\sigma)} \tag{29}$$

where $\mathrm{mix}_\sigma(i,k) \in G$ is defined by

$$[\mathrm{mix}_\sigma(i,k)]_s = \begin{cases} k_s, & s \in \overline{\sigma}, \\ i_s, & s \in \underline{\sigma}. \end{cases}$$

As a result of (28) and (29) we get

$$[BB^*]_{ii} = \sum_{\sigma,\tau\in\{0,1\}^d} \sum_{k\in G} \left| e_{i-k}^{(1,\ldots,1)} \right|^2 \Delta_{\mathrm{mix}_\sigma(i,k)}^{(\sigma)} \overline{\Delta_{\mathrm{mix}_\tau(i,k)}^{(\tau)}}. \tag{30}$$

For a fixed pair (σ,τ), we now partition $\{1,\ldots,d\}$ as follows:

$$\{1,\ldots,d\} = \mathscr{I}_{00} \cup \mathscr{I}_{11} \cup \mathscr{I}_{01} \cup \mathscr{I}_{10},$$

where $\mathscr{I}_{pq} = \{s\in\{1,\ldots,d\} : \sigma_s = p, \tau_s = q\}$ for $p,q \in \{0,1\}$. Let $i^{[pq]}$ stand for the vector containing the $\#\mathscr{I}_{pq}$ entries of i for which the subscript $s \in \mathscr{I}_{pq}$, written in the same order as the d elements of i. Also let $G^{[pq]} = \prod_{s\in\mathscr{I}_{pq}} \mathbb{Z}_{n_s}$. For example, if $d = 4$, $\sigma = (1, 0, 1, 1)$ and $\tau = (0, 0, 1, 0)$, then

$$\mathscr{I}_{10} = \{1, 4\}, \quad i^{[10]} = (i_1, i_4) \quad \text{and} \quad G^{[10]} = \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_4}.$$

Then

$$\left|e_{i-k}^{(1,\dots,1)}\right|^2 = \prod_{p,q=0,1} \left|e_{i^{[pq]}-k^{[pq]}}^{(1,\dots,1)^{[pq]}}\right|^2. \tag{31}$$

Now split the summation over $k \in G$ in (30) additively into four separate summations over $k^{[pq]} \in G^{[pq]}$, i.e.

$$\sum_{k^{[00]}\in G^{[00]}} \sum_{k^{[11]}\in G^{[11]}} \sum_{k^{[01]}\in G^{[01]}} \sum_{k^{[10]}\in G^{[10]}} \prod_{p,q=0,1} \left|e_{i^{[pq]}-k^{[pq]}}^{(1,\dots,1)^{[pq]}}\right|^2 \Delta_{\mathrm{mix}_\sigma(i,k)}^{(\sigma)} \overline{\Delta_{\mathrm{mix}_\tau(i,k)}^{(\tau)}}.$$

As $\mathrm{mix}_\sigma(i, k)^{[00]} = \mathrm{mix}_\tau(i, k)^{[00]} = i^{[00]}$, the unitarity of the circulant matrix $\mathscr{E}_\omega^{(1,\dots,1)}$ involved implies that

$$\sum_{k^{[00]}\in G^{[00]}} \left|e_{i^{[00]}-k^{[00]}}^{(1,\dots,1)^{[00]}}\right|^2 \Delta_{\mathrm{mix}_\sigma(i,k)}^{(\sigma)} \overline{\Delta_{\mathrm{mix}_\tau(i,k)}^{(\tau)}} = \Delta_{\mathrm{mix}_\sigma(i,k)}^{(\sigma)} \overline{\Delta_{\mathrm{mix}_\tau(i,k)}^{(\tau)}}.$$

In other words, the subscripts $s \in \mathscr{I}_{00}$ do not involve any circulant-vector multiplication, just a Schur product between vectors. When summing over $k^{[11]} \in G^{[11]}$, we use that $\mathrm{mix}_\sigma(i, k)^{[11]} = \mathrm{mix}_\tau(i, k)^{[11]} = k^{[11]}$, so that one has to apply a $\#\mathscr{I}_{11}$-index doubly stochastic circulant from the left to the Schur product of two vectors. When summing over $k^{[01]} \in G^{[01]}$, we use that $\mathrm{mix}_\sigma(i, k)^{[01]} = i^{[01]}$ and $\mathrm{mix}_\tau(i, k)^{[01]} = k^{[01]}$, so that one has to apply a doubly stochastic circulant matrix to a vector and then take the Schur product of the resulting vector with another vector. When summing over $k^{[10]} \in G^{[10]}$, we get the complex conjugate of the sum obtained in the $\mathscr{I}_{01}$ case. Thus b_{ii} and $[BB^*]_{ii}$ can be computed in FFT time, i.e., in $\mathrm{O}(n_1 \cdots n_d \log(n_1 \cdots n_d))$ operations, as claimed.

To estimate the computational complexity of the algorithm, we first observe that 2^d d-dimensional FFT's are required to compute the diagonal matrices (26) from the first columns of the σ-circulant matrices of decomposition (12). To compute $\{b_{ii}\}_{i\in G}$ we need $2^d - 1$ circulant-vector products (CVP's), but we can omit them in the total since they appear also in the expression (30) of the denominator of (20).

We now count the total number F_d of CVP's to compute (30) as a function of d. Observe that we only have to compute the contribution to $[BB^*]_{ii}$ for the pair $(\sigma, \tau) \in \{0, 1\}^d \times \{0, 1\}^d$ for which $\sigma \leqslant \tau$ in the lexicographical order on $\{0, 1\}^d$, since the contribution to $[BB^*]_{ii}$ for σ and τ reversed is its complex conjugate. Further, each pair (σ, τ) leads to as many multiindex circulant-vector products as there are sets $\mathscr{I}_{01}$, $\mathscr{I}_{10}$ and $\mathscr{I}_{11}$ nonempty. We thus find $F_1 = 2$, $F_2 = 12$ and $F_3 = 59$ (see Table 1). In general, we have

$$F_d = \frac{3}{2}[4^d - 3^d] + \frac{1}{2}[2^d - 1]. \tag{32}$$

Indeed, writing F_d^{01} and F_d^{11} for the number of circulant-vector products deriving from $\mathscr{I}_{01} \cup \mathscr{I}_{10}$ and $\mathscr{I}_{11}$, so that $F_d = F_d^{01} + F_d^{11}$, we can decompose $\sigma, \tau \in \{0, 1\}^{d+1}$ with $\sigma \leqslant \tau$ as $\sigma = (\sigma_1, \tilde{\sigma})$ and $\tau = (\tau_1, \tilde{\tau})$ $[\tilde{\sigma}, \tilde{\tau} \in \{0, 1\}^d]$. Then we have one of (i) $\sigma_1 = 0$, $\tau_1 = 1$ and $\tilde{\sigma}, \tilde{\tau}$ arbitrary, (ii) $\sigma_1 = \tau_1 = 0$ and $\tilde{\sigma} \leqslant \tilde{\tau}$, or (iii) $\sigma_1 = \tau_1 = 1$ and $\tilde{\sigma} \leqslant \tilde{\tau}$. Thus

$$F_{d+1}^{01} = \underbrace{2^{2d}}_{\sigma_1=0,\tau_1=1} + \underbrace{F_d^{01}}_{\sigma_1=\tau_1=0} + \underbrace{2F_d^{01}}_{\sigma_1=\tau_1=1}, \quad F_1^{01} = 1,$$

Table 1
For $d = 1, 2, 3$ and $\sigma \leqslant \tau$ in $\{0, 1\}^d$ we have indicated how many of the sets $\mathscr{I}_{01}$, $\mathscr{I}_{10}$ and $\mathscr{I}_{11}$, for a given pair (σ, τ) with $\sigma \leqslant \tau$ in the lexicographical order, are nonempty. In that case F_d equals the sum of the integers given in the corresponding table. In fact, $F_1 = 2$, $F_2 = 12$ and $F_3 = 59$

$d = 1$	0	1
0	0	1
1	–	1

$d = 2$	00	01	10	11
00	0	1	1	1
01	–	1	2	2
10	–	–	1	2
11	–	–	–	1

$d = 3$	000	001	010	011	100	101	110	111
000	0	1	1	1	1	1	1	1
001	–	1	2	2	2	2	2	2
010	–	–	1	2	2	2	2	2
011	–	–	–	1	2	3	3	2
100	–	–	–	–	1	2	2	2
101	–	–	–	–	–	1	3	2
110	–	–	–	–	–	–	1	2
111	–	–	–	–	–	–	–	1

$$F_{d+1}^{11} = \underbrace{2F_d^{11} - (2^d - 1)}_{\sigma_1=0 \text{ and } \tau_1=1} + \underbrace{F_d^{11}}_{\sigma_1=\tau_1=0} + \underbrace{2^{d-1}(2^d + 1)}_{\sigma_1=\tau_1=1}, \quad F_1^{11} = 1,$$

where $2^d = \#\{(\tilde{\sigma}, \tilde{\tau}) : \tilde{\sigma} = \tilde{\tau}\}$ and $2^{d-1}(2^d + 1) = \#\{(\tilde{\sigma}, \tilde{\tau}) : \tilde{\sigma} \leqslant \tilde{\tau}\}$. The two difference equations yield $F_d^{01} = 4^d - 3^d$ and $F_d^{11} = \frac{1}{2}[4^d - 3^d + 2^d - 1]$, which imply (32).

To obtain the total number of d-dimensional FFT's (d-FFT) required, we first observe that there are as many nontrivial circulants of the type (31) as there are non empty subsets of $\{1, \ldots, d\}$ (i.e. $2^d - 1$). However, each of these circulant matrices is a tensor product of d one-index circulant matrices, which allows us to compute its eigenvalues by taking products of the eigenvalues of the constituent one-index matrices. This amounts to a reduction of the number of operations involved from $\mathrm{O}(n_1 \cdots n_d \log(n_1 \cdots n_d))$ to $\mathrm{O}(\prod_s n_s + \sum_s n_s \log n_s) = \mathrm{O}(\prod_s n_s)$, where s runs over all indices involved in computing the circulant matrix of type (31), and thus makes this computation negligible in terms of number of d-FFT's when $d > 1$.

Once the eigenvalues are available, each CVP requires 2 d-FFT's. Moreover, 2^d d-FFT's are needed for evaluating the diagonal matrices (26) and 1 to pass from Γ to the preconditioner D. We thus have to perform $2F_d + 2^d + 1$ d-FFT's, which amounts to 8, 29 and 127 for $d = 1, 2, 3$, since one more must be added when $d = 1$ for the computation of the eigenvalues of (31). This number can be further lowered by reusing vectors which appear repeatedly and by exploiting the following symmetry property of the Fourier matrix:

$$\mathscr{F}_\eta(\mathscr{F}_\eta^* \mathbf{x} \circ \mathscr{F}_\eta^* \mathbf{y}) = \mathscr{F}_\eta^*(\mathscr{F}_\eta \mathbf{x} \circ \mathscr{F}_\eta \mathbf{y}),$$

for any $\mathbf{x}, \mathbf{y} \in \mathbb{C}^d$.

5. Tensor products

The special case in which the given Toeplitz matrix A of multiorder n is the tensor product of d Toeplitz matrices of orders $n_1, \ldots, n_d$, respectively, is of particular interest to applications. In this case the matrix A can be written in the form $A = A_1 \otimes \cdots \otimes A_d$, where

$$A_s = (a^{(s)}_{i_s - j_s})_{i_s, j_s = 0}^{n_s - 1}, \quad s = 1, \ldots, d.$$

In this case we have

$$\Delta_i^{(\sigma)} = \prod_{s=1}^{d} \hat{C}_s^{(\sigma_s)}(\omega_s^{-\sigma_s} \eta_s^{-i_s}),$$

where $A_s = C_s^{(0)} + C_s^{(1)}$ is the representation of the one-index Toeplitz matrix A_s as the sum of a circulant and a skew-circulant matrix, as explained in (7). Therefore,

$$b_{ii} = \prod\nolimits_{s=1}^{d} b^{(s)}_{i_s, i_s}, \tag{33}$$

$$[BB^*]_{ii} = \prod_{s=1}^{d} [B_s B_s^*]_{i_s, i_s}, \tag{34}$$

where the factors in the right-hand sides of (33) and (34) are to be computed as in the one-index case. In this situation, the solution of linear system $A\mathbf{x} = \mathbf{b}$, in fact, is itself a one-index problem, since it can be decomposed by solving *in cascade* the $\left(n_1 \cdots n_d \sum_{s=1}^{d} \frac{1}{n_s}\right)$ one-level Toeplitz linear systems

$$\begin{aligned} A_1 \mathbf{y}^{(1)}_{(\bullet, i_2, \ldots, i_d)} &= \mathbf{b}_{(\bullet, i_2, \ldots, i_d)}, \\ A_2 \mathbf{y}^{(2)}_{(i_1, \bullet, \ldots, i_d)} &= \mathbf{y}^{(1)}_{(i_1, \bullet, \ldots, i_d)}, \\ &\vdots \\ A_d \mathbf{x}_{(i_1, i_2, \ldots, \bullet)} &= \mathbf{y}^{(d-1)}_{(i_1, i_2, \ldots, \bullet)}, \end{aligned}$$

where $i_s = 0, \ldots, n_s - 1$, $s = 1, \ldots, d$ and the dot denotes the index used in the matrix product.

When the given Toeplitz matrix A of multiorder n is the sum of N tensor products of one-index Toeplitz matrix, that is

$$A = \sum_{s=1}^{N} (A_{s,1} \otimes \cdots \otimes A_{s,d}),$$

then b_{ii} is the sum of N terms of the form (33), while $[BB^*]_{ii}$ is the sum of N^2 terms of the form (34). This means that the optimal circulant preconditioner of A is the sum of the N optimal circulant preconditioners that correspond to the various tensor product contributions to A, i.e.

$$\mathscr{C}_A = \sum_{s=1}^{N} (\mathscr{C}_{A_{s,1}} \otimes \cdots \otimes \mathscr{C}_{A_{s,d}}).$$

Unfortunately, since the superoptimal circulant preconditioner is computed by evaluating the quotients in the right-hand side of (20), no such result holds for it.

References

[1] R.H. Chan, The spectrum of a family of circulant preconditioned Toeplitz systems, SIAM J. Numer. Anal. 26 (1989) 503–506.
[2] R.H. Chan, X.Q. Jin, M.C. Yeung, The circulant operator in the Banach algebra of matrices, Linear Algebra Appl. 149 (1991) 41–53.
[3] R.H. Chan, M.K. Ng, Conjugate gradient methods for Toeplitz systems, SIAM Rev. 38 (3) (1996) 427–482.
[4] R.H. Chan, G. Strang, Toeplitz equations by conjugate gradients with circulant preconditioner, SIAM J. Sci. Stat. Comput. 10 (1) (1989) 104–119.
[5] T.F. Chan, An optimal circulant preconditioner for Toeplitz systems, SIAM J. Sci. Stat. Comput. 9 (4) (1988) 766–771.
[6] P.J. Davis, Circulant Matrices, John Wiley, New York, 1979.
[7] F. Di Benedetto, C. Estatico, S. Serra Capizzano, Superoptimal preconditioned conjugate gradient iteration for image deblurring, SIAM J. Sci. Comput. 26 (3) (2005) 1012–1035.
[8] M. Hanke, J. Nagy, R. Plemmons, Preconditioned iterative regularization for ill-posed problems, in: Numerical Linear Algebra (Kent, OH, 1992), De Gruyter, Berlin, 1993, pp. 141–163.
[9] G. Strang, A proposal for Toeplitz matrix calculations, Stud. Appl. Math. 74 (1986) 171–176.
[10] V.V. Strela, E.E. Tyrtyshnikov, Which circulant preconditioner is better? Math. Comput. 65 (213) (1996) 137–150.
[11] M. Tismenetsky, A decomposition of Toeplitz matrices and optimal circulant preconditioning, Linear Algebra Appl. 154/156 (1991) 105–121.
[12] E.E. Tyrtyshnikov, Optimal and superoptimal circulant preconditioners, SIAM J. Matrix Anal. Appl. 13 (2) (1992) 459–473.