

## FAST AND NUMERICALLY STABLE ALGORITHMS FOR DISCRETE HARTLEY TRANSFORMS AND APPLICATIONS TO PRECONDITIONING

ANTONIO ARICO\*, STEFANO SERRA-CAPIZZANO†, AND MANFRED TASCHE‡

**Abstract.** The discrete Hartley transforms (DHT) of types I – IV and the related matrix algebras are discussed. We prove that any of these DHTs of length  $N = 2^l$  can be factorized by means of a divide-and-conquer strategy into a product of sparse, orthogonal matrices where in this context sparse means at most two nonzero entries per row and column. The sparsity joint with orthogonality of the matrix factors is the key for proving that these new algorithms have low arithmetic costs equal to  $\frac{5}{2}N \log_2(N) + O(N)$  arithmetic operations and an excellent normwise numerical stability. Further, we consider the optimal Frobenius approximation of a given symmetric Toeplitz matrix generated by an integrable symbol in a Hartley matrix algebra. We give explicit formulas for computing these optimal approximations and discuss the related preconditioned conjugate gradient (PCG) iterations. By using the matrix approximation theory, we prove the strong clustering at unity of the preconditioned matrix sequences under the sole assumption of continuity and positivity of the generating function. The multilevel case is also briefly treated. Some numerical experiments concerning DHT preconditioning are included.

**Key words:** Discrete Hartley transform, discrete Fourier transform, Hartley matrix algebra, fast algorithm, numerical stability, Toeplitz matrix, preconditioning, optimal preconditioner, preconditioned conjugate gradient method, spectral clustering, matrix approximation.

**AMS subject classifications.** 65T50, 65G50, 65F10, 65F15.

**1. Introduction.** Discrete trigonometric transforms are widely used in signal processing, image compression, and numerical linear algebra. Examples of such transforms are discrete Fourier transforms (DFT), discrete cosine transforms, discrete sine transforms, and discrete Hartley transforms (DHT). These transforms are represented by unitary and orthogonal matrices, respectively. The historical roots of the DHT, which is intertwined closely with the DFT, go back to the introduction of the continuous Hartley transform by R. Hartley in 1942. The need to sample signals and approximate the continuous transform on computers led to the DHT. Like the DFT, the DHT is a matrix–vector product, and it requires laborious calculations. The

---

\*Dipartimento di Fisica e Matematica, Università dell’Insubria - Sede di Como, Via Valleggio 11, 22100 Como, Italy; Dipartimento di Matematica “Felice Casorati”, Università di Pavia, Via Ferrata 1, 27100, Pavia, Italy (arico@dimat.unipv.it); the work of this author was partially supported by MIUR, grant number 2002014121.

†Dipartimento di Fisica e Matematica, Università dell’Insubria - Sede di Como, Via Valleggio 11, 22100 Como, Italy (stefano.serrac@uninsubria.it, serra@mail.dm.unipi.it); the work of this author was partially supported by MIUR, grant number 2002014121.

‡Dipartimento di Fisica e Matematica, Università dell’Insubria - Sede di Como, Via Valleggio 11, 22100 Como, Italy (manfred.tasche@uni-rostock.de); this author was Visiting Professor at Insubria University under the MIUR project “Rientro Cervelli” during the year 2004.

key step was performed in 1984 by R.N. Bracewell [6] and the result was the first fast DHT algorithm. This algorithm achieves its speed in the same way as the fast Fourier transform does and computes the DHT of length  $N = 2^t$  in  $O(N \log_2(N))$  operations. Note that every application of the DFT can be realized also by an application of the DHT.

In this paper, we shall concentrate on the construction of real, fast, and recursive DHT algorithms having an excellent numerical stability in floating point arithmetic (see e.g [17, 21]). Further, we will present an application of matrix algebras related to DHT in numerical linear algebra. We will explicitly compute optimal approximations of a symmetric (one-level and multilevel) Toeplitz matrix and we will discuss the related preconditioned conjugate gradient (PCG) iterations. Note that a unified approach to discrete cosine/sine transforms with applications to preconditioning was given by T. Kailath and V. Olshevsky [24, 25] and more general preconditioning results were obtained by Di Benedetto and the second author in [13, 14, 34].

There is a close connection between fast DHT algorithms and factorizations of the corresponding orthogonal Hartley matrix  $H_N \in \mathbf{R}^{N \times N}$  with  $N = 2^t$  ( $t \in \mathbf{N}$ ). Assume that we know a factorization of  $H_N$  into a product of sparse matrices

$$(1.1) \quad H_N = M_N^{(m-1)} \dots M_N^{(0)}, \quad (1 < m \ll N).$$

Then the transformed vector  $H_N \mathbf{x}$  with arbitrary  $\mathbf{x} \in \mathbf{R}^N$  can be computed iteratively by

$$\mathbf{x}^{(s+1)} := M_N^{(s)} \mathbf{x}^{(s)}, \quad \mathbf{x}^{(0)} := \mathbf{x}$$

for  $s = 0, \dots, m-1$  such that  $\mathbf{x}^{(m)} = H_N \mathbf{x}$ . Since all matrix factors in (1.1) are sparse, the arithmetic cost of this method will be low such that the factorization (1.1) of  $H_N$  generates a fast DHT algorithm.

An interesting result in [31] (see also [37]) says that a fast DHT algorithm possesses an excellent numerical stability, if the algorithm is based on a factorization of  $H_N$  into sparse, (almost) orthogonal matrices. Here a matrix is called almost orthogonal, if it is orthogonal up to a positive factor. Therefore, in order to get a real, fast, and numerically stable DHT algorithm, one should be especially interested in a factorization (1.1) with (almost) orthogonal matrix factors of simple structure. Note that various factorizations of  $H_N$  use also non-orthogonal matrix factors.

In this paper, we shall derive fast, completely recursive DHT algorithms of radix-2 length. As usual, all fast DHT algorithms use divide-and-conquer techniques. Further, we shall present complete real factorizations of Hartley matrices into sparse, (almost) orthogonal matrices, where in this context sparse means at most two nonzero entries per row and column. Our DHT algorithms require only permutations, scaling operations, butterfly operations, and plane rotations with small rotation angles.

These matrix factorizations can be also used for a direct (instead of recursive) implementation of the algorithms. The sparsity and orthogonality of the matrix factors in (1.1) are the key for proving that these DHT algorithms have low arithmetic costs and an excellent normwise numerical stability. Using the Wilkinson model for binary floating point arithmetic which is implemented in the IEEE standard, we shall give new, explicit worst case estimates for the errors caused by the applications of our fast DHT algorithms in floating point arithmetic. Further, we consider Hartley matrix algebras which consist of special symmetric Toeplitz–plus–Hankel matrices. Using new representation formulas of these Hartley matrix algebras, we define the best approximation in Frobenius norm of a given symmetric Toeplitz matrix generated by an integrable symbol in the four Hartley matrix algebras. We give explicit formulas for computing these optimal approximations and discuss the related PCG iterations. By using the matrix approximation theory developed in [33, 34], we prove the strong clustering at unity of the preconditioned matrix sequences under the sole assumption of continuity and positivity of the generating function. Concerning the multilevel Toeplitz setting, we carry out the analysis along the same lines as in the one-level case and we show how the new results fit the negative results by the second author and E. Tyrtyshnikov [35, 36].

The paper is organized as follows. In Section 2 we introduce Fourier and Hartley matrices of types I – IV. We sketch the close connection between Fourier and Hartley matrices and discuss the symmetry properties of the Hartley matrices. In this paper, we often use the close relations between Hartley and Fourier matrices. For brevity, we do not discuss the connections between Hartley and cosine/sine matrices. For such a discussion see e.g. [30], pp. 63–66, and for a numerical stability comparison refer to [1]. In Section 3 we recall the main features of circulant and skew-circulant matrices and of their diagonalizations by Fourier matrices of types I – IV. Section 4 is devoted to Hartley matrix algebras. If  $\mathbf{R}^{N \times N}$  is equipped with the Frobenius inner product, then these matrix algebras can be represented as orthogonal sums and corresponding splitting formulas of matrices can be given. In Section 5 we introduce fast recursive DHT algorithms which are based on a recursive factorization of the Hartley matrix of type I of order  $N = 2^t$ . The proof of the factorization is mainly based on divide–and–conquer technique applied directly to a matrix such that the Hartley matrix can be represented as a direct sum of Hartley matrices of half order (and possibly of different type). We also compute the arithmetic costs of these fast DHT algorithms. Corresponding factorizations of Hartley matrices of type I – IV into sparse, (almost) orthogonal matrix factors of simple structure are given. Using these matrix factorizations, we present a comprehensive analysis of the numerical stability of these fast DHT algorithms in Section 6. In Sections 7 and 8, we consider the optimal Frobenius approximation of a given symmetric Toeplitz matrix in a Hartley matrix algebra. With these results, we construct optimal preconditioners for the PCG

method. More precisely, an explicit computation of optimal preconditioners is presented in Subsections 7.1 (for the one-level Toeplitz case) and 8.1 (for the multilevel Toeplitz case), while an approximation analysis for optimal preconditioners is given in Subsections 7.2 (for the one-level case) and 8.2 (for the multilevel case). Subsection 8.3 and Section 9 finish the paper with numerical experiments and conclusions, respectively.

**Contents**

1. Introduction ..... 21  
 2. Fourier and Hartley matrices ..... 24  
 3. Circulant and skew-circulant matrices ..... 28  
 4. Hartley matrix algebras ..... 31  
 5. Fast DHT algorithms ..... 39  
 6. Numerical stability of fast DHT algorithms ..... 48  
 7. Optimal Frobenius approximation of Toeplitz matrices ..... 55  
     7.1. Explicit formulas for optimal preconditioners ..... 56  
     7.2. Approximation results for optimal preconditioners ..... 57  
 8. Multilevel Toeplitz matrices ..... 62  
     8.1. Explicit formulas for optimal preconditioners in the multilevel case ..... 62  
     8.2. Approximation results for optimal preconditioners in the multilevel case ..... 63  
     8.3. Numerical experiments ..... 65  
 9. Conclusions and remarks ..... 66

**2. Fourier and Hartley matrices.** Let  $N \geq 2$  be a given integer. In what follows, we consider *Fourier* and *Hartley matrices of type*  $X \in \{I, II, III, IV\}$  which are defined as

$$F_N^I := \frac{1}{\sqrt{N}} \left( w_N^{jk} \right)_{j,k=0}^{N-1}, \quad F_N^{II} := \frac{1}{\sqrt{N}} \left( w_{2N}^{j(2k+1)} \right)_{j,k=0}^{N-1},$$

$$F_N^{III} := \frac{1}{\sqrt{N}} \left( w_{2N}^{(2j+1)k} \right)_{j,k=0}^{N-1}, \quad F_N^{IV} := \frac{1}{\sqrt{N}} \left( w_{4N}^{(2j+1)(2k+1)} \right)_{j,k=0}^{N-1}$$

and

$$H_N^I := \frac{1}{\sqrt{N}} \left( \text{cas} \left( \frac{2\pi jk}{N} \right) \right)_{j,k=0}^{N-1}, \quad H_N^{II} := \frac{1}{\sqrt{N}} \left( \text{cas} \left( \frac{2\pi j(2k+1)}{2N} \right) \right)_{j,k=0}^{N-1},$$

$$H_N^{III} := \frac{1}{\sqrt{N}} \left( \text{cas} \left( \frac{2\pi (2j+1)k}{2N} \right) \right)_{j,k=0}^{N-1}, \quad H_N^{IV} := \frac{1}{\sqrt{N}} \left( \text{cas} \left( \frac{2\pi (2j+1)(2k+1)}{4N} \right) \right)_{j,k=0}^{N-1},$$

respectively, with  $w_N := \exp(-2\pi\mathbf{i}/N)$  and  $\mathbf{i}^2 = -1$ . Furthermore,  $\text{cas}(x) := \cos(x) + \sin(x)$ , where the word “cas” represents a contraction of the expression “cosine and

sine". For the subsequent analysis it is useful to adopt a more compact notation

$$(2.1) \quad F_N^X := \frac{1}{\sqrt{N}} \left( w_{4N}^{(2j+t^X[1])(2k+t^X[0])} \right)_{j,k=0}^{N-1},$$

$$(2.2) \quad H_N^X := \frac{1}{\sqrt{N}} \left( \cos \left( \frac{2\pi (2j+t^X[1])(2k+t^X[0])}{4N} \right) \right)_{j,k=0}^{N-1},$$

with  $t^I := 0$ ,  $t^{II} := 1$ ,  $t^{III} := 2$ ,  $t^{IV} := 3$ , and where  $t^X[j]$  denotes the  $j$ -th bit of the binary representation of  $t^X$ . We observe that  $F_N^{III} = (F_N^{II})^T$  and  $H_N^{III} = (H_N^{II})^T$  for every  $N \geq 2$ , while for  $N = 2$  we have

$$H_2^I = H_2^{II} = H_2^{III} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = F_2^I, \quad H_2^{IV} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The  $2\pi$ -periodic cas-function is bounded by  $\sqrt{2}$  and fulfills the identities

$$(2.3) \quad \begin{aligned} \text{cas}(x+y) &= \cos(y) \text{cas}(x) + \sin(y) \text{cas}(-x), \\ \text{cas}(x+y) + \text{cas}(x-y) &= 2 \cos(y) \text{cas}(x), \\ \text{cas}(x+\pi) &= -\text{cas}(x). \end{aligned}$$

In our notation, a subscript of a matrix denotes the corresponding order, while a superscript means the "type" of the matrix. The Hartley matrix of type I coincides with the classical one introduced by R.N. Bracewell [6]. Modified Hartley matrices of type II – IV are studied in [22, 19, 7, 15]. Note that Fourier matrices of type I – IV are unitary (it is a trivial check) and that Hartley matrices of type I – IV are orthogonal. This is a consequence of the unitarity of the Fourier matrices (see Lemma 2.1). A *discrete Fourier transform of type  $X \in \{I, II, III, IV\}$  with length  $N$*  (DFT- $X(N)$ ) is the linear mapping which maps any vector  $\mathbf{x} \in \mathbf{C}^N$  into  $F_N^X \mathbf{x}$ . Analogously, a *discrete Hartley transform of type  $X \in \{I, II, III, IV\}$  with length  $N$*  (DHT- $X(N)$ ) is the linear mapping which maps any vector  $\mathbf{x} \in \mathbf{R}^N$  into  $H_N^X \mathbf{x}$ .

First let us agree on some notations and concepts that will be used throughout the paper. By  $I_N$  we denote the identity matrix of order  $N$  and by  $J_N$  the counter-identity matrix which is defined by  $J_N \mathbf{x} = (x_{N-1}, x_{N-2}, \dots, x_0)^T = (x_{N-1-j})_{j=0}^{N-1}$  for every vector  $\mathbf{x} = (x_j)_{j=0}^{N-1}$ . By  $O_N$  we denote the zero matrix of order  $N$ , and by  $\mathbf{o}$  the zero vector. The direct sum of two square matrices  $A, B$  is defined to be the block diagonal matrix  $A \oplus B := \text{diag}(A, B)$ , the Kronecker product of two matrices  $A, B$  is defined to be the block matrix  $A \otimes B := (a_{j,k} B)_{j,k=0}^{N-1}$  with  $A = (a_{j,k})_{j,k=0}^{N-1}$ . Then

$$J'_N := 1 \oplus J_{N-1}, \quad J''_N := (-1) \oplus J_{N-1}$$

are (modified) flip matrices with the obvious property that

$$J'_N \mathbf{x} = (x_0, x_{N-1}, \dots, x_1)^T, \quad J''_N \mathbf{x} = (-x_0, x_{N-1}, \dots, x_1)^T$$

TABLE 2.1

Symmetries, dimensions, and structures of the linear subspaces of even/odd vectors  $\mathbf{x} \in \mathbf{R}^N$ .

symm.	dim.	$N$ even with $d = N/2$	$N$ odd with $d = (N-1)/2$
$J_N$ -even	$\lfloor \frac{N+1}{2} \rfloor$	$(x_0, \dots, x_{d-1}, x_{d-1}, \dots, x_0)^T$	$(x_0, \dots, x_d, x_{d-1}, \dots, x_0)^T$
$J_N$ -odd	$\lfloor \frac{N-1}{2} \rfloor$	$(x_0, \dots, x_{d-1}, -x_{d-1}, \dots, -x_0)^T$	$(x_0, \dots, x_{d-1}, 0, -x_{d-1}, \dots, -x_0)^T$
$J'_N$ -even	$\lceil \frac{N+1}{2} \rceil$	$(x_0, \dots, x_{d-1}, x_d, x_{d-1}, \dots, x_1)^T$	$(x_0, \dots, x_d, x_d, \dots, x_1)^T$
$J'_N$ -odd	$\lceil \frac{N-1}{2} \rceil$	$(0, x_1, \dots, x_{d-1}, 0, -x_{d-1}, \dots, -x_1)^T$	$(0, x_1, \dots, x_d, -x_d, \dots, -x_1)^T$
$J''_N$ -even	$\lceil \frac{N-1}{2} \rceil$	$(0, x_1, \dots, x_{d-1}, x_d, x_{d-1}, \dots, x_1)^T$	$(0, x_1, \dots, x_d, x_d, \dots, x_1)^T$
$J''_N$ -odd	$\lfloor \frac{N+1}{2} \rfloor$	$(x_0, \dots, x_{d-1}, 0, -x_{d-1}, \dots, -x_1)^T$	$(x_0, \dots, x_d, -x_d, \dots, -x_1)^T$

for any vector  $\mathbf{x} = (x_j)_{j=0}^{N-1}$ . Note that  $J_N, J'_N, J''_N$  are symmetric and orthogonal and consequently involution matrices by  $[J_N]^2 = [J'_N]^2 = [J''_N]^2 = I_N$ . A vector  $\mathbf{x} \in \mathbf{C}^N$  is called  $J_N$ -even,  $J'_N$ -even, and  $J''_N$ -even, respectively, if

$$J_N \mathbf{x} = \mathbf{x}, \quad J'_N \mathbf{x} = \mathbf{x}, \quad J''_N \mathbf{x} = \mathbf{x},$$

respectively. Analogously, a vector  $\mathbf{x} \in \mathbf{C}^N$  is called  $J_N$ -odd,  $J'_N$ -odd, and  $J''_N$ -odd, respectively, if

$$J_N \mathbf{x} = -\mathbf{x}, \quad J'_N \mathbf{x} = -\mathbf{x}, \quad J''_N \mathbf{x} = -\mathbf{x},$$

respectively. We use the same definitions of even and odd vectors also for vectors with real entries. The various linear subspaces of even and odd vectors are formed by vectors  $\mathbf{x}$  of the form indicated in Table 2 and, as a consequence, they possess the dimensions displayed in the same table (with respect to the real field  $\mathbf{R}$ ).

The modified Fourier matrices of type  $X \in \{\text{I, II, III, IV}\}$  are connected each other via  $F_N^{\text{I}}$  and more precisely we have

$$(2.4) \quad F_N^{\text{II}} = W_N F_N^{\text{I}}, \quad F_N^{\text{III}} = F_N^{\text{I}} W_N, \quad F_N^{\text{IV}} = w_{4N} W_N F_N^{\text{I}} W_N$$

with

$$W_N := \text{diag} \left( \exp\left(-\frac{\pi j i}{N}\right) \right)_{j=0}^{N-1}.$$

By the definition of Fourier and Hartley matrices given in (2.1) – (2.2), for  $X \in \{\text{I, II, III, IV}\}$ , we infer that

$$H_N^X = \text{Re}(F_N^X) - \text{Im}(F_N^X)$$

with  $\text{Re}(F_N^X) = \frac{1}{2}(F_N^X + \bar{F}_N^X)$  and  $\text{Im}(F_N^X) = \frac{1}{2i}(F_N^X - \bar{F}_N^X)$ . Note that

$$(2.5) \quad \begin{aligned} \bar{F}_N^{\text{I}} &= J'_N F_N^{\text{I}} = F_N^{\text{I}} J'_N, \\ \bar{F}_N^{\text{II}} &= -J''_N F_N^{\text{II}} = F_N^{\text{II}} J_N, \\ \bar{F}_N^{\text{III}} &= J_N F_N^{\text{III}} = -F_N^{\text{III}} J''_N, \\ \bar{F}_N^{\text{IV}} &= -J_N F_N^{\text{IV}} = -F_N^{\text{IV}} J_N. \end{aligned}$$

Then we obtain:

LEMMA 2.1. *Let  $N \geq 2$  be an integer. For  $X \in \{\text{I, II, III, IV}\}$ , it holds*

$$(2.6) \quad H_N^X = F_N^X U_N^X$$

with the unitary matrices

$$U_N^X := \begin{cases} \frac{1}{2}(1 + \mathbf{i})I_N + \frac{1}{2}(1 - \mathbf{i})J'_N & \text{for } X = \text{I,} \\ \frac{1}{2}(1 + \mathbf{i})I_N + \frac{1}{2}(1 - \mathbf{i})J_N & \text{for } X = \text{II,} \\ \frac{1}{2}(1 + \mathbf{i})I_N - \frac{1}{2}(1 - \mathbf{i})J''_N & \text{for } X = \text{III,} \\ \frac{1}{2}(1 + \mathbf{i})I_N - \frac{1}{2}(1 - \mathbf{i})J_N & \text{for } X = \text{IV.} \end{cases}$$

Further  $H_N^X$  is orthogonal.

*Proof.* For simplicity, we consider only the case of  $X = \text{I}$ . By the symmetry of  $J'_N$  we have

$$\begin{aligned} U_N^{\text{I}}[\bar{U}_N^{\text{I}}]^T &= \left[\frac{1}{2}(1 + \mathbf{i})I_N + \frac{1}{2}(1 - \mathbf{i})J'_N\right] \left[\frac{1}{2}(1 - \mathbf{i})I_N + \frac{1}{2}(1 + \mathbf{i})J'_N\right] \\ &= \frac{1}{2}I_N + \frac{1}{4}(1 + \mathbf{i})^2 J'_N + \frac{1}{4}(1 - \mathbf{i})^2 J'_N + \frac{1}{2}I_N = I_N, \end{aligned}$$

and therefore  $U_N^{\text{I}}$  is unitary. Furthermore, by (2.5),  $H_N^{\text{I}}$  can be represented in the form

$$\begin{aligned} H_N^{\text{I}} &= \text{Re}(F_N^{\text{I}}) - \text{Im}(F_N^{\text{I}}) \\ &= \frac{1}{2}(F_N^{\text{I}} + F_N^{\text{I}}J'_N) - \frac{1}{2\mathbf{i}}(F_N^{\text{I}} - F_N^{\text{I}}J'_N) \\ &= F_N^{\text{I}} \left[\frac{1}{2}(1 + \mathbf{i})I_N + \frac{1}{2}(1 - \mathbf{i})J'_N\right] = F_N^{\text{I}} U_N^{\text{I}}. \end{aligned}$$

Hence the real matrix  $H_N^{\text{I}}$  is the product of unitary matrices and therefore  $H_N^{\text{I}}$  is orthogonal.  $\square$

As a consequence of (2.5) – (2.6), we deduce the following relationships:

$$(2.7) \quad \begin{aligned} H_N^{\text{I}} J'_N &= J'_N H_N^{\text{I}}, & H_N^{\text{II}} J_N &= -J''_N H_N^{\text{II}}, \\ H_N^{\text{III}} J''_N &= -J_N H_N^{\text{III}}, & H_N^{\text{IV}} J_N &= J_N H_N^{\text{IV}}. \end{aligned}$$

From (2.7) we can easily deduce several symmetry properties of DHTs. Note that Lemma 2.2 improves a corresponding result of [19, 20].

LEMMA 2.2. *Let  $N \geq 2$  be an integer. Then we have the following properties:*

1. DHT-I( $N$ ) maps any  $J'_N$ -even vector  $\mathbf{x} \in \mathbf{R}^N$  into a  $J'_N$ -even vector  $H_N^{\text{I}}\mathbf{x} = F_N^{\text{I}}\mathbf{x}$  and any  $J'_N$ -odd vector  $\mathbf{y} \in \mathbf{R}^N$  into a  $J'_N$ -odd vector  $H_N^{\text{I}}\mathbf{y} = \mathbf{i}F_N^{\text{I}}\mathbf{y}$ .
2. DHT-II( $N$ ) maps any  $J_N$ -even vector  $\mathbf{x} \in \mathbf{R}^N$  into a  $J''_N$ -odd vector  $H_N^{\text{II}}\mathbf{x} = F_N^{\text{II}}\mathbf{x}$  and any  $J_N$ -odd vector  $\mathbf{y} \in \mathbf{R}^N$  into a  $J''_N$ -even vector  $H_N^{\text{II}}\mathbf{y} = \mathbf{i}F_N^{\text{II}}\mathbf{y}$ .

3. DHT-III( $N$ ) maps any  $J_N''$ -even vector  $\mathbf{x} \in \mathbf{R}^N$  into a  $J_N$ -odd vector  $H_N^{\text{III}}\mathbf{x} = \mathbf{i}F_N^{\text{III}}\mathbf{x}$  and any  $J_N''$ -odd vector  $\mathbf{y} \in \mathbf{R}^N$  into a  $J_N$ -even vector  $H_N^{\text{III}}\mathbf{y} = F_N^{\text{III}}\mathbf{y}$ .
4. DHT-IV( $N$ ) maps any  $J_N$ -even vector  $\mathbf{x} \in \mathbf{R}^N$  into a  $J_N$ -even vector  $H_N^{\text{IV}}\mathbf{x} = \mathbf{i}F_N^{\text{IV}}\mathbf{x}$  and any  $J_N$ -odd vector  $\mathbf{y} \in \mathbf{R}^N$  into a  $J_N$ -odd vector  $H_N^{\text{IV}}\mathbf{y} = F_N^{\text{IV}}\mathbf{y}$ .

*Proof.* The results follow immediately from (2.5) – (2.7). For simplicity, we sketch the proof only in the case of  $X = \text{II}$ . Let  $\mathbf{x} \in \mathbf{R}^N$  be  $J_N$ -even. Then  $H_N^{\text{II}}\mathbf{x}$  is  $J_N''$ -odd, since a direct computation, based on (2.7), shows that  $H_N^{\text{II}}\mathbf{x} = H_N^{\text{II}}J_N\mathbf{x} = -J_N''(H_N^{\text{II}}\mathbf{x})$ . If  $\mathbf{y} \in \mathbf{R}^N$  is  $J_N$ -odd, then  $H_N^{\text{II}}\mathbf{y}$  is  $J_N''$ -even because (again by (2.7))  $H_N^{\text{II}}\mathbf{y} = H_N^{\text{II}}(-J_N\mathbf{y}) = J_N''(H_N^{\text{II}}\mathbf{y})$ . Moreover, by (2.5) – (2.6), for  $J_N$ -even  $\mathbf{x} \in \mathbf{R}^N$ , we obtain that

$$H_N^{\text{II}}\mathbf{x} = F_N^{\text{II}}U_N^{\text{II}}\mathbf{x} = F_N^{\text{II}}\left(\frac{1+\mathbf{i}}{2}I_N + \frac{1-\mathbf{i}}{2}J_N\right)\mathbf{x} = F_N^{\text{II}}\left(\frac{1+\mathbf{i}}{2} + \frac{1-\mathbf{i}}{2}\right)\mathbf{x} = F_N^{\text{II}}\mathbf{x}.$$

In a similar manner, we can prove the relation  $H_N^{\text{II}}\mathbf{y} = \mathbf{i}F_N^{\text{II}}\mathbf{y}$  for any  $J_N$ -odd vector  $\mathbf{y} \in \mathbf{R}^N$ .  $\square$

**3. Circulant and skew-circulant matrices.** The *circulant matrix* of  $\mathbf{x} = (x_j)_{j=0}^{N-1} \in \mathbf{C}^N$  is defined by

$$\text{circ}(\mathbf{x}) := (x_{(j-k)\bmod N})_{j,k=0}^{N-1}$$

(see [11], p. 66). Here  $(j-k)\bmod N$  denotes the nonnegative residue of  $j-k$  modulo  $N$  for  $j, k = 0, \dots, N-1$ , i.e.

$$(j-k)\bmod N = \begin{cases} j-k & \text{if } j-k \geq 0, \\ j-k+N & \text{if } j-k < 0. \end{cases}$$

Note that for  $\mathbf{e}_0 = (1, 0, \dots, 0)^T \in \mathbf{C}^N$ , we have  $\text{circ}(\mathbf{e}_0) = I_N$ . The circulant matrices are related to the cyclic convolution. More precisely for  $\mathbf{x}, \mathbf{y} \in \mathbf{C}^N$ , their *cyclic convolution* is defined as

$$\mathbf{z} = (z_j)_{j=0}^{N-1} = \mathbf{x} * \mathbf{y} := \text{circ}(\mathbf{x})\mathbf{y}$$

i.e.

$$z_j = \sum_{k=0}^{N-1} x_{(j-k)\bmod N} y_k.$$

The cyclic convolution is an associative, commutative and distributive multiplication with the unit element  $\mathbf{e}_0$ . By the cyclic convolution property of DFT-I( $N$ ), for all  $\mathbf{x}, \mathbf{y} \in \mathbf{C}^N$  we have

$$[F_N^{\text{I}}]^\alpha(\mathbf{x} * \mathbf{y}) = \sqrt{N} ([F_N^{\text{I}}]^\alpha\mathbf{x}) \circ ([F_N^{\text{I}}]^\alpha\mathbf{y}) \quad (\alpha \in \{-1, 1\}),$$

where  $\circ$  denotes the componentwise product of vectors:  $\mathbf{x} \circ \mathbf{y} := \text{diag}(\mathbf{x}) \mathbf{y}$ . From the above relation with  $\alpha = -1$  we deduce that

$$[F_N^I]^{-1} \text{circ}(\mathbf{x}) \mathbf{y} = \text{diag}\left(\sqrt{N} [F_N^I]^{-1} \mathbf{x}\right) [F_N^I]^{-1} \mathbf{y}$$

and hence

$$(3.1) \quad \text{circ}(\mathbf{x}) = F_N^I \text{diag}\left(\sqrt{N} [F_N^I]^{-1} \mathbf{x}\right) [F_N^I]^{-1}.$$

As a consequence, any circulant matrix of order  $N$  is diagonalized by  $F_N^I$ .

Let  $\text{Circ}(N)$  be the set of all circulant matrices of order  $N$ . By (3.1), it follows that  $\text{Circ}(N)$  is a commutative algebra. Let  $\text{Diag}(F_N^X)$  with  $X \in \{\text{I, II, III, IV}\}$  denote the set of matrices  $A_N$  such that  $[F_N^X]^{-1} A_N F_N^X$  is a diagonal matrix. It is clear that  $\text{Diag}(F_N^X)$  is a commutative algebra as well.

LEMMA 3.1. *Let  $N \geq 2$  be an integer. Then we have*

$$\text{Diag}(F_N^I) = \text{Diag}(F_N^{\text{III}}) = \text{Circ}(N).$$

Moreover, for every  $\mathbf{x} \in \mathbf{C}^N$

$$(3.2) \quad [F_N^I]^{-1} \text{circ}(\mathbf{x}) F_N^I = \text{diag}(\sqrt{N} [F_N^I]^{-1} \mathbf{x}),$$

$$(3.3) \quad [F_N^{\text{III}}]^{-1} \text{circ}(\mathbf{x}) F_N^{\text{III}} = \text{diag}(\sqrt{N} [F_N^I]^{-1} \mathbf{x}).$$

*Proof.* Formula (3.2) follows directly from (3.1). By (2.4) we know that  $F_N^{\text{III}} = F_N^I W_N$  with unitary diagonal matrix  $W_N$ . Therefore the columns of  $F_N^{\text{III}}$  are just multiple (via coefficients of modulus 1) of the columns of  $F_N^I$  and consequently (3.3) is valid, since both the set of columns of  $F_N^I, F_N^{\text{III}}$  are systems of eigenvectors of  $\text{circ}(\mathbf{x})$  related to the same ordering of the eigenvalues.  $\square$

The *skew-circulant matrix* of  $\mathbf{x} = (x_j)_{j=0}^{N-1} \in \mathbf{C}^N$  is defined by

$$\text{scirc}(\mathbf{x}) := \left( (\text{sign}(j-k)) x_{(j-k) \bmod N} \right)_{j,k=0}^{N-1},$$

with  $\text{sign}(j-k) := 1$  if  $j-k \geq 0$  and  $\text{sign}(j-k) := -1$  otherwise (see [11], p. 83). Note that  $\text{scirc}(\mathbf{e}_0) = I_N$ . The skew-circulant matrices are closely related to the skew-cyclic convolution. More in detail, for  $\mathbf{x}, \mathbf{y} \in \mathbf{C}^N$ , the *skew-cyclic convolution* is defined as

$$\mathbf{z} = (z_j)_{j=0}^{N-1} = \mathbf{x} \diamond \mathbf{y} := \text{scirc}(\mathbf{x}) \mathbf{y}$$

i.e.

$$z_j = \sum_{k=0}^{N-1} (\text{sign}(j-k)) x_{(j-k) \bmod N} y_k.$$

The skew-cyclic convolution is also an associative, commutative and distributive multiplication with the unit element  $\mathbf{e}_0$ . By DFT-III( $N$ ), the skew-cyclic convolution can be transferred into a componentwise product. This corresponds to the so-called skew-cyclic convolution property of DFT-III( $N$ ):

LEMMA 3.2. *For all  $\mathbf{x}, \mathbf{y} \in \mathbf{C}^N$ , we have*

$$F_N^{\text{III}}(\mathbf{x} \diamond \mathbf{y}) = \sqrt{N} (F_N^{\text{III}} \mathbf{x}) \circ (F_N^{\text{III}} \mathbf{y}).$$

*Proof.* Let  $\mathbf{z} = (z_j)_{j=0}^{N-1} = \mathbf{x} \diamond \mathbf{y}$  and let us compute the  $s$ -th component of  $F_N^{\text{III}} \mathbf{z}$ :

$$(F_N^{\text{III}} \mathbf{z})_s = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} w_{2N}^{(2s+1)j} z_j \quad (s = 0, \dots, N-1)$$

with

$$z_j = \sum_{k=0}^{N-1} (\text{sign}(j-k)) x_{(j-k) \bmod N} y_k.$$

Thus by a change in the summation order, we get

$$(F_N^{\text{III}} \mathbf{z})_s = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} y_k w_{2N}^{(2s+1)k} \left( \sum_{j=0}^{N-1} w_{2N}^{(2s+1)(j-k)} (\text{sign}(j-k)) x_{(j-k) \bmod N} \right).$$

Now we substitute  $t := (j-k) \bmod N$  in the second sum. For arbitrary fixed  $k \in \{0, \dots, N-1\}$ , the new index  $t$  runs over the whole set  $\{0, \dots, N-1\}$  as  $j$  moves over it and moreover

$$(\text{sign}(j-k)) w_{2N}^{(2s+1)(j-k)} = w_{2N}^{(2s+1)t},$$

since  $w_{2N}^N = -1$ . Hence we obtain

$$\begin{aligned} (F_N^{\text{III}} \mathbf{z})_s &= \frac{1}{\sqrt{N}} \left( \sum_{k=0}^{N-1} y_k w_{2N}^{(2s+1)k} \right) \left( \sum_{t=0}^{N-1} x_t w_{2N}^{(2s+1)t} \right) \\ &= \sqrt{N} (F_N^{\text{III}} \mathbf{y})_s (F_N^{\text{III}} \mathbf{x})_s \quad (s = 0, \dots, N-1) \end{aligned}$$

and the proof is complete.  $\square$

Since  $J_N F_N^{\text{III}} = [F_N^{\text{II}}]^{-1}$ , and  $J_N \text{diag}(\mathbf{v}) = \text{diag}(J_N \mathbf{v}) J_N$ , by Lemma 3.2 it follows that

$$[F_N^{\text{II}}]^{-1} \text{scirc}(\mathbf{x}) \mathbf{y} = \text{diag}(\sqrt{N} [F_N^{\text{II}}]^{-1} \mathbf{x}) [F_N^{\text{II}}]^{-1} \mathbf{y}$$

and consequently

$$(3.4) \quad \text{scirc}(\mathbf{x}) = F_N^{\text{II}} \text{diag}(\sqrt{N} [F_N^{\text{II}}]^{-1} \mathbf{x}) [F_N^{\text{II}}]^{-1}.$$

Thus any skew-circulant matrix of order  $N$  is diagonalized by  $F_N^{\text{II}}$ . Let  $\text{Scirc}(N)$  be the set of all skew-circulant matrices of order  $N$ . By (3.4), we infer that  $\text{Scirc}(N)$  is a commutative algebra.

LEMMA 3.3. *Let  $N \geq 2$  be an integer. Then we have*

$$\text{Diag}(F_N^{\text{II}}) = \text{Diag}(F_N^{\text{IV}}) = \text{Scirc}(N).$$

Moreover, for every  $\mathbf{x} \in \mathbf{C}^N$

$$(3.5) \quad [F_N^{\text{II}}]^{-1} \text{scirc}(\mathbf{x}) F_N^{\text{II}} = \text{diag}\left(\sqrt{N} [F_N^{\text{II}}]^{-1} \mathbf{x}\right),$$

$$(3.6) \quad [F_N^{\text{IV}}]^{-1} \text{scirc}(\mathbf{x}) F_N^{\text{IV}} = \text{diag}\left(\sqrt{N} [F_N^{\text{II}}]^{-1} \mathbf{x}\right).$$

*Proof.* Formula (3.5) is a plain consequence of (3.4). By (2.4) we know that  $F_N^{\text{IV}} = w_{4N} F_N^{\text{II}} W_N$  with diagonal unitary matrix  $W_N$ . Therefore the columns of  $F_N^{\text{IV}}$  are just multiples (with factors of modulus 1) of the columns of  $F_N^{\text{II}}$  and consequently (3.6) is valid, since both the sets of columns are systems of eigenvectors of  $\text{scirc}(\mathbf{x})$  related to the same ordering of the eigenvalues.  $\square$

By  $\dim_{\mathbf{F}}(\cdot)$  we denote the dimension of a linear space with respect to the field  $\mathbf{F} \in \{\mathbf{R}, \mathbf{C}\}$ . Note that  $\dim_{\mathbf{C}}(\text{Circ}(N)) = \dim_{\mathbf{C}}(\text{Scirc}(N)) = N$  and  $\dim_{\mathbf{R}}(\text{Circ}(N)) = \dim_{\mathbf{R}}(\text{Scirc}(N)) = 2N$ . By  $\text{Circ}_0(N)$  and  $\text{Circ}_1(N)$ , we denote the set of all circulant matrices  $\text{circ}(\mathbf{x})$  with  $J_N$ -even and  $J_N$ -odd vectors  $\mathbf{x} \in \mathbf{R}^N$ , respectively. Analogously, by  $\text{Scirc}_0(N)$  and  $\text{Scirc}_1(N)$ , we denote the set of all skew-circulant matrices  $\text{scirc}(\mathbf{x})$  with  $J_N''$ -even and  $J_N''$ -odd vectors  $\mathbf{x} \in \mathbf{R}^N$ , respectively. Obviously,  $\text{Circ}_0(N)$ ,  $\text{Circ}_1(N)$ ,  $\text{Scirc}_0(N)$  and  $\text{Scirc}_1(N)$  are linear subspaces of  $\mathbf{R}^{N \times N}$  and, in view of Table 2, we have

$$\begin{aligned} \dim_{\mathbf{R}}(\text{Circ}_0(N)) &= \lceil (N+1)/2 \rceil, & \dim_{\mathbf{R}}(\text{Circ}_1(N)) &= \lfloor (N-1)/2 \rfloor, \\ \dim_{\mathbf{R}}(\text{Scirc}_0(N)) &= \lceil (N-1)/2 \rceil, & \dim_{\mathbf{R}}(\text{Scirc}_1(N)) &= \lfloor (N+1)/2 \rfloor. \end{aligned}$$

The sum  $\dim_{\mathbf{R}}(\text{Circ}_0(N)) + \dim_{\mathbf{R}}(\text{Circ}_1(N))$  is  $N$ , as the dimension of the real subspace of  $\text{Circ}(N)$  of real circulant matrices. The same is true for  $\text{Scirc}_0(N)$  and  $\text{Scirc}_1(N)$  of course. Later we will see that  $\text{Circ}_0(N)$  and  $\text{Scirc}_1(N)$  are commutative matrix algebras, while  $\text{Circ}_1(N)$  and  $\text{Scirc}_0(N)$  do not possess such a structure.

**4. Hartley matrix algebras.** By  $\text{Diag}(H_N^X)$  with  $X \in \{\text{I, II, III, IV}\}$  we denote the set of real matrices  $A_N$  such that  $[H_N^X]^{-1} A_N H_N^X$  is a diagonal matrix. It is clear that  $\text{Diag}(H_N^X)$  is the (real) commutative matrix algebra related to  $\text{DHT-}X(N)$  which is called *Hartley matrix algebra of type  $X$* . We will see that for every  $X \in \{\text{I, II, III, IV}\}$ , the sets  $\text{Diag}(H_N^X)$  consist of special symmetric Toeplitz-plus-Hankel matrices. More specifically, in this section we improve previous results by Bini and Favati [4] and by Bortoletti and Di Fiore [5]. Bini and Favati [4] have considered

the Hartley matrix algebra of type I and, recently, Bortoletti and Di Fiore [5] have analyzed the Hartley matrix algebras of types  $X \in \{\text{II}, \text{III}, \text{IV}\}$ .

In the space  $\mathbf{R}^{N \times N}$ , we introduce the *Frobenius inner product*

$$(4.1) \quad \langle A_N, B_N \rangle := \text{tr}(A_N^T B_N), \quad (A_N, B_N \in \mathbf{R}^{N \times N})$$

with  $\text{tr}(A)$  denoting the trace of a square matrix  $A$ . Then  $\mathbf{R}^{N \times N}$  with the latter inner product is a Hilbert space. By  $\oplus$  we denote the orthogonal sum of linear subspaces of  $\mathbf{R}^{N \times N}$ . The following theorem improves a corresponding result of Bini and Favati [4], more specifically, the orthogonal sum representation (4.2) and formula (4.3) are new.

**THEOREM 4.1.** *Let  $N \geq 2$  be an integer. Then the following orthogonal sum representation is valid*

$$(4.2) \quad \text{Diag}(H_N^I) = \text{Circ}_0(N) \oplus J'_N \text{Circ}_1(N).$$

Moreover, for arbitrary  $\mathbf{x} \in \mathbf{R}^N$

$$(4.3) \quad H_N^I \text{diag}(\sqrt{N}\mathbf{x}) H_N^I = \text{circ}(H_N^I \mathbf{x}_+) - J'_N \text{circ}(H_N^I \mathbf{x}_-)$$

with  $J'_N$ -even and  $J'_N$ -odd parts of  $\mathbf{x}$  defined by

$$\mathbf{x}_+ := \frac{1}{2}(\mathbf{x} + J'_N \mathbf{x}), \quad \mathbf{x}_- := \frac{1}{2}(\mathbf{x} - J'_N \mathbf{x}).$$

*Proof.* 1. For arbitrary  $\mathbf{x} \in \mathbf{R}^N$ , let  $X_N^I := H_N^I \text{diag}(\sqrt{N}\mathbf{x}) H_N^I$ . Since  $H_N^I$  is symmetric and orthogonal, we have  $X_N^I = H_N^I \text{diag}(\sqrt{N}\mathbf{x}) [H_N^I]^{-1}$ . By (2.6) we know that

$$\begin{aligned} H_N^I &= F_N^I \left( \frac{1}{2}(1 + \mathbf{i})I_N + \frac{1}{2}(1 - \mathbf{i})J'_N \right), \\ [H_N^I]^{-1} &= \left( \frac{1}{2}(1 - \mathbf{i})I_N + \frac{1}{2}(1 + \mathbf{i})J'_N \right) [F_N^I]^{-1}. \end{aligned}$$

Hence we obtain

$$X_N^I = F_N^I \sqrt{N} \left( \frac{1}{2} \text{diag}(\mathbf{x}) - \frac{\mathbf{i}}{2} J'_N \text{diag}(\mathbf{x}) + \frac{\mathbf{i}}{2} \text{diag}(\mathbf{x}) J'_N + \frac{1}{2} J'_N \text{diag}(\mathbf{x}) J'_N \right) [F_N^I]^{-1}.$$

By

$$(4.4) \quad J'_N \text{diag}(\mathbf{x}) J'_N = \text{diag}(J'_N \mathbf{x}), \quad \text{diag}(\mathbf{x}) J'_N = J'_N \text{diag}(J'_N \mathbf{x})$$

it follows that  $X_N^I = F_N^I \sqrt{N} (\text{diag}(\mathbf{x}_+) - \mathbf{i} J'_N \text{diag}(\mathbf{x}_-)) [F_N^I]^{-1}$ . By (3.2), we get

$$X_N^I = \text{circ}(F_N^I \mathbf{x}_+) - \mathbf{i} J'_N \text{circ}(F_N^I \mathbf{x}_-).$$

Since  $\mathbf{x} \in \mathbf{R}^N$ , we deduce by Lemma 2.2 that

$$F_N^I \mathbf{x}_+ = H_N^I \mathbf{x}_+, \quad F_N^I \mathbf{x}_- = -\mathbf{i} H_N^I \mathbf{x}_-,$$

where  $H_N^I \mathbf{x}_+$  is  $J'_N$ -even and  $H_N^I \mathbf{x}_-$  is  $J'_N$ -odd. Thus (4.3) is proven.

2. By (4.3) it follows that

$$\text{Diag}(H_N^I) \subseteq \text{Circ}_0(N) + J'_N \text{Circ}_1(N),$$

since  $H_N^I \mathbf{x}_+ \in \mathbf{R}^N$  is  $J'_N$ -even and  $H_N^I \mathbf{x}_- \in \mathbf{R}^N$  is  $J'_N$ -odd by Lemma 2.2.

Conversely, if  $\text{circ}(\mathbf{a}) - J'_N \text{circ}(\mathbf{b})$  is an arbitrary matrix belonging to  $\text{Circ}_0(N) + J'_N \text{Circ}_1(N)$  (so that  $\mathbf{a} \in \mathbf{R}^N$  is  $J'_N$ -even and  $\mathbf{b} \in \mathbf{R}^N$  is  $J'_N$ -odd), then we form  $\mathbf{x} := H_N^I(\mathbf{a} + \mathbf{b}) \in \mathbf{R}^N$ . By Lemma 2.2, the  $J'_N$ -even and  $J'_N$ -odd parts of  $\mathbf{x}$  read as follows

$$\mathbf{x}_+ = \frac{1}{2}(\mathbf{x} + J'_N \mathbf{x}) = H_N^I \mathbf{a}, \quad \mathbf{x}_- = \frac{1}{2}(\mathbf{x} - J'_N \mathbf{x}) = H_N^I \mathbf{b}.$$

Since  $[H_N^I]^{-1} = H_N^I$ , we get  $\mathbf{a} = H_N^I \mathbf{x}_+$  and  $\mathbf{b} = H_N^I \mathbf{x}_-$ . Using (4.3), we see that

$$\begin{aligned} \text{circ}(\mathbf{a}) - J'_N \text{circ}(\mathbf{b}) &= \text{circ}(H_N^I \mathbf{x}_+) - J'_N \text{circ}(H_N^I \mathbf{x}_-) \\ &= H_N^I \text{diag}(\sqrt{N} \mathbf{x}) H_N^I \in \text{Diag}(H_N^I). \end{aligned}$$

Thus we also have

$$\text{Circ}_0(N) + J'_N \text{Circ}_1(N) \subseteq \text{Diag}(H_N^I).$$

3. Finally we show that (4.2) is an orthogonal sum representation of  $\text{Diag}(H_N^I)$ . For arbitrary  $J'_N$ -even  $\mathbf{a} \in \mathbf{R}^N$  and  $J'_N$ -odd  $\mathbf{b} \in \mathbf{R}^N$ , we have to prove that the inner product

$$\langle \text{circ}(\mathbf{a}), J'_N \text{circ}(\mathbf{b}) \rangle = \text{tr} [(\text{circ}(\mathbf{a}))^T J'_N \text{circ}(\mathbf{b})]$$

vanishes. By (4.3), we have

$$\text{circ}(\mathbf{a}) = H_N^I \text{diag}(\sqrt{N} \hat{\mathbf{a}}) H_N^I, \quad J'_N \text{circ}(\mathbf{b}) = -H_N^I \text{diag}(\sqrt{N} \hat{\mathbf{b}}) H_N^I$$

with  $\hat{\mathbf{a}} = (\hat{a}_j)_{j=0}^{N-1} = H_N^I \mathbf{a}$  and  $\hat{\mathbf{b}} = (\hat{b}_j)_{j=0}^{N-1} = H_N^I \mathbf{b}$ . By the trace theorem, it follows that

$$\langle \text{circ}(\mathbf{a}), J'_N \text{circ}(\mathbf{b}) \rangle = -N \text{tr} \left[ H_N^I \text{diag}(\hat{\mathbf{a}}) \text{diag}(\hat{\mathbf{b}}) H_N^I \right] = -N \sum_{j=0}^{N-1} \hat{a}_j \hat{b}_j.$$

But by Lemma 2.2,  $\hat{\mathbf{a}}$  is  $J'_N$ -even and  $\hat{\mathbf{b}}$  is  $J'_N$ -odd and therefore

$$\hat{a}_j = \hat{a}_{N-j}, \quad \hat{b}_j = -\hat{b}_{N-j}, \quad \hat{b}_0 = 0 \quad (j = 1, \dots, N-1).$$

Then we conclude that

$$\langle \text{circ}(\mathbf{a}), J'_N \text{circ}(\mathbf{b}) \rangle = -N (\hat{a}_0 \hat{b}_0 + \frac{1}{2} \sum_{j=1}^{N-1} (\hat{a}_j \hat{b}_j + \hat{a}_{N-j} \hat{b}_{N-j})) = 0.$$

This completes the proof.  $\square$

EXAMPLE 4.2. Let  $\mathbf{x} \in \mathbf{R}^N$  be an arbitrary vector with  $J'_N$ -even and  $J'_N$ -odd parts indicated by  $\mathbf{x}_+$  and  $\mathbf{x}_-$ , respectively. For the sake of notational simplicity, we introduce the vectors  $\mathbf{a} = H_N^I \mathbf{x}_+$  and  $\mathbf{b} = H_N^I \mathbf{x}_-$  which, by Lemma 2.2, are  $J'_N$ -even and  $J'_N$ -odd, respectively. For  $N = 4$  the latter means  $\mathbf{a} = (a_0, a_1, a_2, a_1)^T$  and  $\mathbf{b} = (0, b_1, 0, -b_1)^T$  with arbitrary  $a_0, a_1, a_2, b_1 \in \mathbf{R}$  and

$$X_4^I = \begin{pmatrix} a_0 & a_1 & a_2 & a_1 \\ a_1 & a_0 & a_1 & a_2 \\ a_2 & a_1 & a_0 & a_1 \\ a_1 & a_2 & a_1 & a_0 \end{pmatrix} + \begin{pmatrix} 0 & b_1 & 0 & -b_1 \\ b_1 & 0 & -b_1 & 0 \\ 0 & -b_1 & 0 & b_1 \\ -b_1 & 0 & b_1 & 0 \end{pmatrix}$$

with  $X_4^I = 2 H_4^I \text{diag}(\mathbf{x}) H_4^I$ . For  $N = 5$ , following the same lines we have  $\mathbf{a} = (a_0, a_1, a_2, a_2, a_1)^T$  and  $\mathbf{b} = (0, b_1, b_2, -b_2, -b_1)^T$  with arbitrary  $a_0, a_1, a_2, b_1, b_2 \in \mathbf{R}$  and

$$X_5^I = \begin{pmatrix} a_0 & a_1 & a_2 & a_2 & a_1 \\ a_1 & a_0 & a_1 & a_2 & a_2 \\ a_2 & a_1 & a_0 & a_1 & a_2 \\ a_2 & a_2 & a_1 & a_0 & a_1 \\ a_1 & a_2 & a_2 & a_1 & a_0 \end{pmatrix} + \begin{pmatrix} 0 & b_1 & b_2 & -b_2 & -b_1 \\ b_1 & b_2 & -b_2 & -b_1 & 0 \\ b_2 & -b_2 & -b_1 & 0 & b_1 \\ -b_2 & -b_1 & 0 & b_1 & b_2 \\ -b_1 & 0 & b_1 & b_2 & -b_2 \end{pmatrix}$$

with  $X_5^I = \sqrt{5} H_5^I \text{diag}(\mathbf{x}) H_5^I$ . We just observe that  $\text{Diag}(H_N^I)$  consists of special symmetric Toeplitz-plus-Hankel matrices.

Now we improve some results concerning  $\text{Diag}(H_N^X)$  with  $X \in \{\text{II}, \text{III}, \text{IV}\}$ , recently obtained by Bortoletti and Di Fiore [5]. More in detail, the orthogonal sum representations of  $\text{Diag}(H_N^X)$ ,  $X \in \{\text{II}, \text{III}, \text{IV}\}$ , and formulae (4.6), (4.9), and (4.11) are new.

THEOREM 4.3. *Let  $N \geq 2$  be an integer. Then the following orthogonal sum representation is valid*

$$(4.5) \quad \text{Diag}(H_N^{\text{II}}) = \text{Scirc}_1(N) \oplus J_N'' \text{Scirc}_0(N).$$

Moreover, for arbitrary  $\mathbf{x} \in \mathbf{R}^N$

$$(4.6) \quad H_N^{\text{II}} \text{diag}(\sqrt{N} \mathbf{x}) H_N^{\text{III}} = \text{scirc}(H_N^{\text{II}} \mathbf{x}_+) + J_N'' \text{scirc}(H_N^{\text{II}} \mathbf{x}_-)$$

with  $J_N$ -even and  $J_N$ -odd parts of  $\mathbf{x}$  defined by

$$\mathbf{x}_+ := \frac{1}{2}(\mathbf{x} + J_N \mathbf{x}), \quad \mathbf{x}_- := \frac{1}{2}(\mathbf{x} - J_N \mathbf{x}).$$

*Proof.* 1. For arbitrary  $\mathbf{x} \in \mathbf{R}^N$ , let  $X_N^{\text{II}} := H_N^{\text{II}} \text{diag}(\sqrt{N} \mathbf{x}) H_N^{\text{III}}$ . Since  $H_N^{\text{II}}$  is orthogonal and its transpose coincides with  $H_N^{\text{III}}$ , it holds

$$X_N^{\text{II}} = H_N^{\text{II}} \text{diag}(\sqrt{N} \mathbf{x}) [H_N^{\text{II}}]^{-1}.$$

By (2.6) we know that

$$\begin{aligned} H_N^{\text{II}} &= F_N^{\text{II}} \left( \frac{1}{2}(1 + \mathbf{i})I_N + \frac{1}{2}(1 - \mathbf{i})J_N \right), \\ [H_N^{\text{II}}]^{-1} &= \left( \frac{1}{2}(1 - \mathbf{i})I_N + \frac{1}{2}(1 + \mathbf{i})J_N \right) [F_N^{\text{II}}]^{-1}. \end{aligned}$$

Hence we obtain

$$X_N^{\text{II}} = F_N^{\text{II}} \frac{\sqrt{N}}{2} [\text{diag}(\mathbf{x}) - \mathbf{i} J_N \text{diag}(\mathbf{x}) + \mathbf{i} \text{diag}(\mathbf{x}) J_N + J_N \text{diag}(\mathbf{x}) J_N] [F_N^{\text{II}}]^{-1}.$$

By

$$(4.7) \quad J_N \text{diag}(\mathbf{x}) J_N = \text{diag}(J_N \mathbf{x}), \quad \text{diag}(\mathbf{x}) J_N = J_N \text{diag}(J_N \mathbf{x})$$

it follows that  $X_N^{\text{II}} = F_N^{\text{II}} \sqrt{N} [\text{diag}(\mathbf{x}_+) - \mathbf{i} J_N \text{diag}(\mathbf{x}_-)] [F_N^{\text{II}}]^{-1}$ . We have by (2.5) that

$$F_N^{\text{II}} J_N = -J_N'' F_N^{\text{II}}.$$

Thus we get by (3.5) that

$$X_N^{\text{II}} = \text{scirc}(F_N^{\text{II}} \mathbf{x}_+) + \mathbf{i} J_N'' \text{scirc}(F_N^{\text{II}} \mathbf{x}_-).$$

For arbitrary  $\mathbf{x} \in \mathbf{R}^N$ , we get by Lemma 2.2 that

$$F_N^{\text{II}} \mathbf{x}_+ = H_N^{\text{II}} \mathbf{x}_+, \quad F_N^{\text{II}} \mathbf{x}_- = -\mathbf{i} H_N^{\text{II}} \mathbf{x}_-,$$

where  $H_N^{\text{II}} \mathbf{x}_+$  is  $J_N''$ -odd and  $H_N^{\text{II}} \mathbf{x}_-$  is  $J_N''$ -even. Thus (4.6) is proven.

2. By (4.6) it follows that

$$\text{Diag}(H_N^{\text{II}}) \subseteq \text{Scirc}_1(N) + J_N'' \text{Scirc}_0(N)$$

since  $H_N^{\text{II}} \mathbf{x}_+ \in \mathbf{R}^N$  is  $J_N''$ -odd and  $H_N^{\text{II}} \mathbf{x}_- \in \mathbf{R}^N$  is  $J_N''$ -even by Lemma 2.2.

Conversely, if  $\text{scirc}(\mathbf{a}) + J_N'' \text{scirc}(\mathbf{b})$  is an arbitrary matrix belonging to  $\text{Scirc}_1(N) + J_N'' \text{Scirc}_0(N)$  (so that  $\mathbf{a} \in \mathbf{R}^N$  is  $J_N''$ -odd and  $\mathbf{b} \in \mathbf{R}^N$  is  $J_N''$ -even), then we form  $\mathbf{x} := H_N^{\text{III}}(\mathbf{a} + \mathbf{b}) \in \mathbf{R}^N$ . By Lemma 2.2, the  $J_N$ -even and  $J_N$ -odd parts of  $\mathbf{x}$  read as follows

$$\mathbf{x}_+ = \frac{1}{2}(\mathbf{x} + J_N \mathbf{x}) = H_N^{\text{III}} \mathbf{a}, \quad \mathbf{x}_- = \frac{1}{2}(\mathbf{x} - J_N \mathbf{x}) = H_N^{\text{III}} \mathbf{b}.$$

Since  $[H_N^{\text{III}}]^{-1} = H_N^{\text{II}}$ , we get  $\mathbf{a} = H_N^{\text{II}} \mathbf{x}_+$  and  $\mathbf{b} = H_N^{\text{II}} \mathbf{x}_-$ . Using (4.6), we see that

$$\begin{aligned} \text{scirc}(\mathbf{a}) + J_N'' \text{scirc}(\mathbf{b}) &= \text{scirc}(H_N^{\text{II}} \mathbf{x}_+) + J_N'' \text{scirc}(H_N^{\text{II}} \mathbf{x}_-) \\ &= H_N^{\text{II}} \text{diag}(\sqrt{N} \mathbf{x}) H_N^{\text{III}} \in \text{Diag}(H_N^{\text{II}}). \end{aligned}$$

Thus we also have

$$\text{Scirc}_1(N) + J_N'' \text{Scirc}_0(N) \subseteq \text{Diag}(H_N^{\text{II}}).$$

3. Finally we show that (4.5) is an orthogonal sum representation of  $\text{Diag}(H_N^{\text{II}})$ . For arbitrary  $J_N''$ -odd  $\mathbf{a} \in \mathbf{R}^N$  and  $J_N''$ -even  $\mathbf{b} \in \mathbf{R}^N$ , we have to prove that the inner product

$$\langle \text{scirc}(\mathbf{a}), J_N'' \text{circ}(\mathbf{b}) \rangle = \text{tr} [(\text{scirc}(\mathbf{a}))^T J_N'' \text{scirc}(\mathbf{b})]$$

vanishes. By (4.6), we have

$$\text{scirc}(\mathbf{a}) = H_N^{\text{II}} \text{diag}(\sqrt{N} \hat{\mathbf{a}}) H_N^{\text{III}}, \quad J_N'' \text{scirc}(\mathbf{b}) = H_N^{\text{II}} \text{diag}(\sqrt{N} \hat{\mathbf{b}}) H_N^{\text{III}}$$

with  $\hat{\mathbf{a}} = (\hat{a}_j)_{j=0}^{N-1} = H_N^{\text{III}} \mathbf{a}$  and  $\hat{\mathbf{b}} = (\hat{b}_j)_{j=0}^{N-1} = H_N^{\text{III}} \mathbf{b}$ . By the trace theorem, it follows that

$$\langle \text{scirc}(\mathbf{b}), J_N'' \text{scirc}(\mathbf{a}) \rangle = N \text{tr} \left[ H_N^{\text{II}} \text{diag}(\hat{\mathbf{a}}) \text{diag}(\hat{\mathbf{b}}) H_N^{\text{III}} \right] = N \sum_{j=0}^{N-1} \hat{a}_j \hat{b}_j.$$

But by Lemma 2.2,  $\hat{\mathbf{a}}$  is  $J_N$ -odd and  $\hat{\mathbf{b}}$  is  $J_N$ -even and therefore

$$\hat{a}_j = -\hat{a}_{N-1-j}, \quad \hat{b}_j = \hat{b}_{N-1-j}, \quad (j = 0, \dots, N-1).$$

Then we conclude that

$$\langle \text{scirc}(\mathbf{b}), J_N'' \text{scirc}(\mathbf{a}) \rangle = \frac{N}{2} \sum_{j=0}^{N-1} (\hat{a}_j \hat{b}_j + \hat{a}_{N-1-j} \hat{b}_{N-1-j}) = 0$$

and this completes the proof.  $\square$

EXAMPLE 4.4. Let  $\mathbf{x} \in \mathbf{R}^N$  be an arbitrary vector with  $J_N$ -even and  $J_N$ -odd parts denoted by  $\mathbf{x}_+$  and  $\mathbf{x}_-$ , respectively. We define the vectors  $\mathbf{a} = H_N^{\text{II}} \mathbf{x}_+$  and  $\mathbf{b} = H_N^{\text{II}} \mathbf{x}_-$  which, by Lemma 2.2, are  $J_N''$ -odd and  $J_N''$ -even, respectively. For  $N = 4$  the latter means  $\mathbf{a} = (a_0, a_1, 0, -a_1)^T$  and  $\mathbf{b} = (0, b_1, b_2, b_1)^T$  with arbitrary real numbers  $a_0, a_1, b_1, b_2$  and

$$X_4^{\text{II}} = \begin{pmatrix} a_0 & a_1 & 0 & -a_1 \\ a_1 & a_0 & a_1 & 0 \\ 0 & a_1 & a_0 & a_1 \\ -a_1 & 0 & a_1 & a_0 \end{pmatrix} + \begin{pmatrix} 0 & b_1 & b_2 & b_1 \\ b_1 & b_2 & b_1 & 0 \\ b_2 & b_1 & 0 & -b_1 \\ b_1 & 0 & -b_1 & -b_2 \end{pmatrix}$$

with  $X_4^{\text{II}} = 2 H_4^{\text{II}} \text{diag}(\mathbf{x}) H_4^{\text{III}}$ .

For  $N = 5$ , following the same lines we have  $\mathbf{a} = (a_0, a_1, a_2, -a_2, -a_1)^T$  and  $\mathbf{b} = (0, b_1, b_2, b_2, b_1)^T$  with arbitrary real numbers  $a_0, a_1, a_2, b_1, b_2$  and

$$X_5^{\text{II}} = \begin{pmatrix} a_0 & a_1 & a_2 & -a_2 & -a_1 \\ a_1 & a_0 & a_1 & a_2 & -a_2 \\ a_2 & a_1 & a_0 & a_1 & a_2 \\ -a_2 & a_2 & a_1 & a_0 & a_1 \\ -a_1 & -a_2 & a_2 & a_1 & a_0 \end{pmatrix} + \begin{pmatrix} 0 & b_1 & b_2 & b_2 & b_1 \\ b_1 & b_2 & b_2 & b_1 & 0 \\ b_2 & b_2 & b_1 & 0 & -b_1 \\ b_2 & b_1 & 0 & -b_1 & -b_2 \\ b_1 & 0 & -b_1 & -b_2 & -b_2 \end{pmatrix}$$

with  $X_5^{\text{II}} = \sqrt{5} H_5^{\text{II}} \text{diag}(\mathbf{x}) H_5^{\text{II}}$ . We notice that  $\text{Diag}(H_N^{\text{II}})$  consists of special symmetric Toeplitz–plus–Hankel matrices.

**THEOREM 4.5.** *Let  $N \geq 2$  be an integer. Then the following orthogonal sum representation is valid*

$$(4.8) \quad \text{Diag}(H_N^{\text{II}}) = \text{Circ}_0(N) \oplus J_N \text{Circ}_1(N).$$

Moreover, for arbitrary  $\mathbf{x} \in \mathbf{R}^N$

$$(4.9) \quad H_N^{\text{II}} \text{diag}(\sqrt{N}\mathbf{x}) H_N^{\text{II}} = \text{circ}(H_N^{\text{I}}\mathbf{x}_+) - J_N \text{circ}(H_N^{\text{I}}\mathbf{x}_-)$$

with  $\mathbf{x}_+$ ,  $\mathbf{x}_-$  being the  $J_N$ -even and  $J_N$ -odd parts of  $\mathbf{x}$ , respectively.

The proof is totally similar to the one of Theorem 4.3 and is omitted here.

**EXAMPLE 4.6.** We use the same notations as in Example 4.2. For  $N = 4$ , we get

$$X_4^{\text{III}} = \begin{pmatrix} a_0 & a_1 & a_2 & a_1 \\ a_1 & a_0 & a_1 & a_2 \\ a_2 & a_1 & a_0 & a_1 \\ a_1 & a_2 & a_1 & a_0 \end{pmatrix} + \begin{pmatrix} b_1 & 0 & -b_1 & 0 \\ 0 & -b_1 & 0 & b_1 \\ -b_1 & 0 & b_1 & 0 \\ 0 & b_1 & 0 & -b_1 \end{pmatrix}$$

with  $X_4^{\text{III}} = 2 H_4^{\text{III}} \text{diag}(\mathbf{x}) H_4^{\text{III}}$ . In the case  $N = 5$ , we obtain

$$X_5^{\text{III}} = \begin{pmatrix} a_0 & a_1 & a_2 & a_2 & a_1 \\ a_1 & a_0 & a_1 & a_2 & a_2 \\ a_2 & a_1 & a_0 & a_1 & a_2 \\ a_2 & a_2 & a_1 & a_0 & a_1 \\ a_1 & a_2 & a_2 & a_1 & a_0 \end{pmatrix} + \begin{pmatrix} b_1 & b_2 & -b_2 & -b_1 & 0 \\ b_2 & -b_2 & -b_1 & 0 & b_1 \\ -b_2 & -b_1 & 0 & b_1 & b_2 \\ -b_1 & 0 & b_1 & b_2 & -b_2 \\ 0 & b_1 & b_2 & -b_2 & -b_1 \end{pmatrix}$$

with  $X_5^{\text{III}} = \sqrt{5} H_5^{\text{III}} \text{diag}(\mathbf{x}) H_5^{\text{III}}$ . Also for type III we find that  $\text{Diag}(H_N^{\text{III}})$  consists of special symmetric Toeplitz–plus–Hankel matrices. Moreover note that  $\text{Diag}(H_N^{\text{III}}) \neq \text{Diag}(H_N^{\text{I}})$ .

**THEOREM 4.7.** *Let  $N \geq 2$  be an integer. Then the following orthogonal sum representation is valid*

$$(4.10) \quad \text{Diag}(H_N^{\text{IV}}) = \text{Scirc}_1(N) \oplus J_N \text{Scirc}_0(N).$$

Moreover, for arbitrary  $\mathbf{x} \in \mathbf{R}^N$

$$(4.11) \quad H_N^{\text{IV}} \text{diag}(\sqrt{N}\mathbf{x}) H_N^{\text{IV}} = \text{scirc}(H_N^{\text{II}}\mathbf{x}_+) + J_N \text{scirc}(H_N^{\text{II}}\mathbf{x}_-)$$

with  $\mathbf{x}_+$ ,  $\mathbf{x}_-$  being the  $J_N$ -even and  $J_N$ -odd parts of  $\mathbf{x}$ , respectively.

The proof follows similar lines as the proof of Theorem 4.3 and is omitted here.

EXAMPLE 4.8. We use the same notations as in Example 4.4. For  $N = 4$  we get

$$X_4^{\text{IV}} = \begin{pmatrix} a_0 & a_1 & 0 & -a_1 \\ a_1 & a_0 & a_1 & 0 \\ 0 & a_1 & a_0 & a_1 \\ -a_1 & 0 & a_1 & a_0 \end{pmatrix} + \begin{pmatrix} b_1 & b_2 & b_1 & 0 \\ b_2 & b_1 & 0 & -b_1 \\ b_1 & 0 & -b_1 & -b_2 \\ 0 & -b_1 & -b_2 & -b_1 \end{pmatrix}$$

with  $X_4^{\text{IV}} = 2 H_4^{\text{IV}} \text{diag}(\mathbf{x}) H_4^{\text{IV}}$ , while for  $N = 5$ , we obtain

$$X_5^{\text{IV}} = \begin{pmatrix} a_0 & a_1 & a_2 & -a_2 & -a_1 \\ a_1 & a_0 & a_1 & a_2 & -a_2 \\ a_2 & a_1 & a_0 & a_1 & a_2 \\ -a_2 & a_2 & a_1 & a_0 & a_1 \\ -a_1 & -a_2 & a_2 & a_1 & a_0 \end{pmatrix} + \begin{pmatrix} b_1 & b_2 & b_2 & b_1 & 0 \\ b_2 & b_2 & b_1 & 0 & -b_1 \\ b_2 & b_1 & 0 & -b_1 & -b_2 \\ b_1 & 0 & -b_1 & -b_2 & -b_2 \\ 0 & -b_1 & -b_2 & -b_2 & -b_1 \end{pmatrix}$$

with  $X_5^{\text{IV}} = \sqrt{5} H_5^{\text{IV}} \text{diag}(\mathbf{x}) H_5^{\text{IV}}$ . We observe that  $\text{Diag}(H_N^{\text{IV}})$  consists of special symmetric Toeplitz–plus–Hankel matrices. Finally notice that all the Hartley matrix algebras  $\text{Diag}(H_N^X)$  of types  $X \in \{\text{I}, \text{II}, \text{III}, \text{IV}\}$  are pairwise distinct.

REMARK 4.9. The sets  $\text{Circ}_0(N)$ ,  $\text{Circ}_1(N)$ ,  $\text{Scirc}_0(N)$ , and  $\text{Scirc}_1(N)$  are linear spaces. Furthermore  $\text{Circ}_0(N)$  and  $\text{Scirc}_1(N)$  are commutative matrix algebras, but  $\text{Circ}_1(N)$  and  $\text{Scirc}_0(N)$  do not possess such a structure.

If  $\text{circ}(\mathbf{a}), \text{circ}(\mathbf{b}) \in \text{Circ}_0(N)$  with  $J'_N$ -even vectors  $\mathbf{a}, \mathbf{b} \in \mathbf{R}^N$ , then by Theorem 4.1 and Lemma 2.2 we get

$$\text{circ}(\mathbf{a}) = H_N^{\text{I}} \text{diag}(\sqrt{N} \hat{\mathbf{a}}) H_N^{\text{I}}, \quad \text{circ}(\mathbf{b}) = H_N^{\text{I}} \text{diag}(\sqrt{N} \hat{\mathbf{b}}) H_N^{\text{I}}$$

with  $J'_N$ -even vectors  $\hat{\mathbf{a}} = H_N^{\text{I}} \mathbf{a}$  and  $\hat{\mathbf{b}} = H_N^{\text{I}} \mathbf{b}$  and hence

$$\begin{aligned} \text{circ}(\mathbf{a}) \text{circ}(\mathbf{b}) &= H_N^{\text{I}} \text{diag}(\sqrt{N} \hat{\mathbf{a}}) \text{diag}(\sqrt{N} \hat{\mathbf{b}}) H_N^{\text{I}} \\ &= H_N^{\text{I}} \text{diag}(N \hat{\mathbf{a}} \circ \hat{\mathbf{b}}) H_N^{\text{I}} \in \text{Circ}_0(N), \end{aligned}$$

since the componentwise product  $\hat{\mathbf{a}} \circ \hat{\mathbf{b}} \in \mathbf{R}^N$  remains  $J'_N$ -even. Analogously, by Theorem 4.7 and Lemma 2.2 it can be seen that  $\text{Scirc}_1(N)$  is a matrix algebra.

On the other hand, if  $\text{circ}(\mathbf{a}), \text{circ}(\mathbf{b}) \in \text{Circ}_1(N)$  with  $J'_N$ -odd vectors  $\mathbf{a}, \mathbf{b} \in \mathbf{R}^N$ , then again by Theorem 4.1 and Lemma 2.2 we obtain

$$\text{circ}(\mathbf{a}) = -J'_N H_N^{\text{I}} \text{diag}(\sqrt{N} \hat{\mathbf{a}}) H_N^{\text{I}}, \quad \text{circ}(\mathbf{b}) = -J'_N H_N^{\text{I}} \text{diag}(\sqrt{N} \hat{\mathbf{b}}) H_N^{\text{I}}$$

with  $J'_N$ -odd vectors  $\hat{\mathbf{a}} = H_N^{\text{I}} \mathbf{a}$  and  $\hat{\mathbf{b}} = H_N^{\text{I}} \mathbf{b}$ . Therefore by (2.7) we deduce

$$\begin{aligned} \text{circ}(\mathbf{a}) \text{circ}(\mathbf{b}) &= -J'_N H_N^{\text{I}} \text{diag}(\sqrt{N} \hat{\mathbf{a}}) H_N^{\text{I}} (-J'_N) H_N^{\text{I}} \text{diag}(\sqrt{N} \hat{\mathbf{b}}) H_N^{\text{I}} \\ &= H_N^{\text{I}} J'_N \text{diag}(\sqrt{N} \hat{\mathbf{a}}) H_N^{\text{I}} H_N^{\text{I}} J'_N \text{diag}(\sqrt{N} \hat{\mathbf{b}}) H_N^{\text{I}} \\ &= H_N^{\text{I}} J'_N \text{diag}(\sqrt{N} \hat{\mathbf{a}}) J'_N \text{diag}(\sqrt{N} \hat{\mathbf{b}}) H_N^{\text{I}} \\ &= H_N^{\text{I}} \text{diag}(\sqrt{N} J'_N \hat{\mathbf{a}}) \text{diag}(\sqrt{N} \hat{\mathbf{b}}) H_N^{\text{I}} \\ &= -H_N^{\text{I}} \text{diag}(\sqrt{N} \hat{\mathbf{a}}) \text{diag}(\sqrt{N} \hat{\mathbf{b}}) H_N^{\text{I}} \\ &= -H_N^{\text{I}} \text{diag}(N \hat{\mathbf{a}} \circ \hat{\mathbf{b}}) H_N^{\text{I}} \in \text{Circ}_0(N), \end{aligned}$$

since  $\hat{\mathbf{a}} \circ \hat{\mathbf{b}}$  is now  $J'_N$ -even. Thus  $\text{Circ}_1(N)$  is not a matrix algebra. In a similar way, by Theorem 4.7 and Lemma 2.2 we can prove that  $\text{Scirc}_0(N)$  is not a matrix algebra.

REMARK 4.10. We observe that the proofs in parts 3 of Theorems 4.1, 4.3, 4.5, and 4.7 regarding the orthogonality of the decompositions (4.2), (4.5), (4.8), and (4.10), and the proofs of the statements reported in Remark 4.9 only concern  $\text{Circ}_0(N)$ ,  $\text{Circ}_1(N)$ ,  $\text{Scirc}_0(N)$ , and  $\text{Scirc}_1(N)$ , i.e. real linear subspaces of the circulant and skew-circulant matrices. However the employed tools include formulae (4.3), (4.6), (4.9), and (4.11), and Lemma 2.2 which basically concerns links between Hartley matrices (and Hartley matrix algebras) and Fourier matrices (and circulant, skew-circulant matrices). Of course, we can also prove the above statements only referring to circulant, skew-circulant, and Fourier matrices. The (modest) price to pay is that we have to work in the complex field.

As an example, let us prove the orthogonality of the decompositions (4.2) reported in part 3 of the proof of Theorem 4.1 by using purely ‘‘circulant tools’’. For arbitrary  $J'_N$ -even  $\mathbf{a} \in \mathbf{R}^N$  and  $J'_N$ -odd  $\mathbf{b} \in \mathbf{R}^N$ , we have to prove that the inner product

$$\langle \text{circ}(\mathbf{a}), J'_N \text{circ}(\mathbf{b}) \rangle = \text{tr} [(\text{circ}(\mathbf{a}))^T J'_N \text{circ}(\mathbf{b})]$$

vanishes. Setting  $\hat{\mathbf{a}} = (\hat{a}_j)_{j=0}^{N-1} = [F_N^I]^{-1} \mathbf{a}$  and  $\hat{\mathbf{b}} = (\hat{b}_j)_{j=0}^{N-1} = [F_N^I]^{-1} \mathbf{b}$ , by (3.2), we have

$$\text{circ}(\mathbf{a}) = F_N^I \text{diag}(\sqrt{N}\hat{\mathbf{a}}) [F_N^I]^{-1}, \quad \text{circ}(\mathbf{b}) = F_N^I \text{diag}(\sqrt{N}\hat{\mathbf{b}}) [F_N^I]^{-1}.$$

By the trace theorem and by (2.5), it follows that

$$\begin{aligned} \langle \text{circ}(\mathbf{a}), J'_N \text{circ}(\mathbf{b}) \rangle &= N \text{tr} \left[ F_N^I \text{diag}(\hat{\mathbf{a}}) [F_N^I]^{-1} J'_N F_N^I \text{diag}(\hat{\mathbf{b}}) [F_N^I]^{-1} \right] \\ &= N \text{tr} \left[ F_N^I \text{diag}(\hat{\mathbf{a}}) J'_N \text{diag}(\hat{\mathbf{b}}) [F_N^I]^{-1} \right] \\ &= N \times \begin{cases} \hat{a}_0 \hat{b}_0 & \text{if } N \text{ is odd,} \\ (\hat{a}_0 \hat{b}_0 + \hat{a}_{N/2} \hat{b}_{N/2}) & \text{otherwise.} \end{cases} \end{aligned}$$

By direct check,  $\hat{\mathbf{b}} \in \mathbf{C}^N$  is  $J'_N$ -odd and therefore  $\hat{b}_0 = 0$  and  $\hat{b}_{N/2} = 0$ , if  $N$  is even. Then we conclude that  $\langle \text{circ}(\mathbf{a}), J'_N \text{circ}(\mathbf{b}) \rangle = 0$ .

REMARK 4.11. The results contained in Theorems 4.1, 4.3, 4.5, and 4.7 are (partially) summarized in Table 4.11. More precisely,  $\mathbf{x}_+$  indicates the even vector and  $\mathbf{x}_-$  the odd vector obtained from  $\mathbf{x}$  with respect to the matrix  $J'_N$  for the Hartley matrix algebras of types I and III and with respect to  $J_N$  for the Hartley matrix algebras of types II and IV.

**5. Fast DHT algorithms.** The aim of the present section is to introduce and discuss fast DHT algorithms (cf. [2, 7, 5]). We start with some preliminaries. For

TABLE 4.1  
Symmetries of vectors and structures of Hartley matrix algebras.

$X$	$X_N^X := H_N^X \operatorname{diag}(\sqrt{N} \mathbf{x}) [H_N^X]^{-1} \in \operatorname{Diag}(H_N^X)$	$\mathbf{x}_\pm$	structure
I	$X_N^I = \operatorname{circ}(H_N^I \mathbf{x}_+) - J'_N \operatorname{circ}(H_N^I \mathbf{x}_-)$	$J'_N$	$\operatorname{circ}(J'_N\text{-even}) - J'_N \operatorname{circ}(J'_N\text{-odd})$
II	$X_N^{II} = \operatorname{scirc}(H_N^{II} \mathbf{x}_+) + J_N'' \operatorname{scirc}(H_N^{II} \mathbf{x}_-)$	$J_N$	$\operatorname{scirc}(J_N''\text{-odd}) + J_N'' \operatorname{scirc}(J_N''\text{-even})$
III	$X_N^{III} = \operatorname{circ}(H_N^I \mathbf{x}_+) - J_N \operatorname{circ}(H_N^I \mathbf{x}_-)$	$J'_N$	$\operatorname{circ}(J'_N\text{-even}) - J_N \operatorname{circ}(J'_N\text{-odd})$
IV	$X_N^{IV} = \operatorname{scirc}(H_N^{II} \mathbf{x}_+) + J_N \operatorname{scirc}(H_N^{II} \mathbf{x}_-)$	$J_N$	$\operatorname{scirc}(J_N''\text{-odd}) + J_N \operatorname{scirc}(J_N''\text{-even})$

$N \geq 4$ ,  $P_N$  denotes the *even-odd permutation matrix* (or *2-stride permutation matrix*) defined by

$$P_N \mathbf{x} := \begin{cases} (x_0, x_2, \dots, x_{N-2}, x_1, x_3, \dots, x_{N-1})^T & \text{for even } N, \\ (x_0, x_2, \dots, x_{N-1}, x_1, x_3, \dots, x_{N-2})^T & \text{for odd } N \end{cases}$$

with  $\mathbf{x} = (x_j)_{j=0}^{N-1} \in \mathbf{R}^N$ . Note that  $P_N^{-1} = P_N^T$  is the  $\lfloor N/2 \rfloor$ -stride permutation matrix (or perfect shuffle matrix). First we recall intertwining relations of Hartley matrices. The following lemma improves a corresponding result of [19].

LEMMA 5.1. *Let  $N \geq 2$  be an integer. The Hartley matrices (2.2) satisfy the following intertwining relations*

$$(5.1) \quad \begin{aligned} H_N^{II} &= T'_N H_N^I = H_N^{IV} T_N, & H_N^{III} &= H_N^I T'_N = T_N H_N^{IV}, \\ H_N^{IV} &= T'_N H_N^I T_N = T_N H_N^I T'_N = H_N^{II} T_N = T_N H_N^{III} \end{aligned}$$

with the symmetric orthogonal twiddle matrices

$$\begin{aligned} T'_N &:= \operatorname{diag} \left( \cos \left( \frac{\pi j}{N} \right) \right)_{j=0}^{N-1} + \left[ \operatorname{diag} \left( \sin \left( \frac{\pi j}{N} \right) \right)_{j=0}^{N-1} \right] J'_N, \\ T_N &:= \operatorname{diag} \left( \cos \left( \frac{(2j+1)\pi}{2N} \right) \right)_{j=0}^{N-1} + \left[ \operatorname{diag} \left( \sin \left( \frac{(2j+1)\pi}{2N} \right) \right)_{j=0}^{N-1} \right] J_N. \end{aligned}$$

*Proof.* Obviously, we have  $T'_N = (T'_N)^T$  and  $T_N = (T_N)^T$ , since  $(\sin(\pi j/N))_{j=0}^{N-1}$  is  $J'_N$ -even and  $(\sin((2j+1)\pi/(2N)))_{j=0}^{N-1}$  is  $J_N$ -even. Finally by direct computation we see that

$$T'_N (T'_N)^T = \operatorname{diag} \left( \cos^2 \left( \frac{\pi j}{N} \right) \right)_{j=0}^{N-1} + \operatorname{diag} \left( \sin^2 \left( \frac{\pi j}{N} \right) \right)_{j=0}^{N-1} = I_N,$$

since

$$\begin{aligned} & \left[ \operatorname{diag} \left( \cos \left( \frac{\pi j}{N} \right) \right)_{j=0}^{N-1} \right] J'_N \left[ \operatorname{diag} \left( \sin \left( \frac{\pi j}{N} \right) \right)_{j=0}^{N-1} \right] \\ & + \left[ \operatorname{diag} \left( \sin \left( \frac{\pi j}{N} \right) \right)_{j=0}^{N-1} \right] J'_N \left[ \operatorname{diag} \left( \cos \left( \frac{\pi j}{N} \right) \right)_{j=0}^{N-1} \right] \\ & = \left[ 0 \oplus \operatorname{diag} \left( \sin \left( \frac{\pi j}{N} \right) \cos \left( \frac{\pi(N-j)}{N} \right) + \sin \left( \frac{\pi(N-j)}{N} \right) \cos \left( \frac{\pi j}{N} \right) \right)_{j=1}^{N-1} \right] J'_N \end{aligned}$$

and by well-known trigonometric formulae,

$$\begin{aligned} & \sin\left(\frac{\pi j}{N}\right) \cos\left(\frac{\pi(N-j)}{N}\right) + \sin\left(\frac{\pi(N-j)}{N}\right) \cos\left(\frac{\pi j}{N}\right) = \\ & \sin\left(\frac{\pi j}{N} + \frac{\pi(N-j)}{N}\right) = \sin(\pi) = 0, \quad (j = 1, \dots, N-1). \end{aligned}$$

In a similar way, we have  $T_N (T_N)^T = I_N$ . Using (2.2) and the basic properties of the cas-function (2.3), we obtain  $H_N^{\text{II}} = T'_N H_N^{\text{I}}$  and  $H_N^{\text{IV}} = T_N H_N^{\text{III}}$ . Hence by transposing, we have  $H_N^{\text{III}} = H_N^{\text{I}} T'_N$ ,  $H_N^{\text{IV}} = H_N^{\text{II}} T_N$  and this completes the proof.  $\square$

REMARK 5.2. The matrices  $T_N$  and  $T'_N$  are *cross-shaped*, since their non-vanishing entries lie on a cross. The central entry of the cross of  $T'_N$  with even  $N$  and of  $T_N$  with odd  $N$  is equal to 1. In the other cases, this central element simply does not exist. After suitable permutations, the matrices  $T'_N$  and  $T_N$  are direct sums of 1 and of rotation-reflection matrices of order 2 (with angles  $\pi j/N$  ( $j = 1, \dots, \lfloor (N-1)/2 \rfloor$ ) and  $(2j+1)\pi/(2N)$  ( $j = 0, \dots, \lfloor (N-2)/2 \rfloor$ ), respectively). Therefore, the matrices  $T_N$  and  $T'_N$  are called *twiddle matrices*.

In the following we use Lemma 5.1 for the construction of fast and numerically stable algorithms for DHT- $X(N)$  with  $X \in \{\text{II}, \text{III}, \text{IV}\}$ . Fast algorithms which use pre- and postprocessing (with multiplication of  $T'_N$  and  $T_N$ , respectively) are designed only for DHT-I( $N$ ). Let us recall an orthogonal factorization of  $H_N^{\text{I}}$  which holds for even  $N \geq 4$  and is similar to the one presented in [5].

LEMMA 5.3. *Let  $N \geq 4$  be an even integer. The Hartley matrix  $H_N^{\text{I}}$  can be factored in the form*

$$(5.2) \quad H_N^{\text{I}} = P_N^T (I_2 \otimes H_{N/2}^{\text{I}}) (I_{N/2} \oplus T'_{N/2}) (H_2^{\text{I}} \otimes I_{N/2}).$$

*Proof.* We show (5.2) by using a divide-and-conquer technique. First we permute the rows of  $H_N^{\text{I}}$  according to the permutation matrix  $P_N$  and we write the result as a block matrix

$$P_N H_N^{\text{I}} = \frac{1}{\sqrt{N}} \begin{pmatrix} \left( \left( \cos\left(\frac{4\pi j k}{N}\right) \right)_{j,k=0}^{N/2-1} & \left( \cos\left(\frac{4\pi j(N/2+k)}{N}\right) \right)_{j,k=0}^{N/2-1} \right. \\ \left. \left( \cos\left(\frac{2\pi(2j+1)k}{N}\right) \right)_{j,k=0}^{N/2-1} & \left( \cos\left(\frac{2\pi(2j+1)(N/2+k)}{N}\right) \right)_{j,k=0}^{N/2-1} \right) \end{pmatrix}.$$

By using (2.2), (2.3), and (5.1) we write the latter in a more informative way:

$$\begin{aligned}
P_N H_N^I &= \frac{1}{\sqrt{2}} \begin{pmatrix} H_{N/2}^I & H_{N/2}^I \\ H_{N/2}^{III} & -H_{N/2}^{III} \end{pmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{pmatrix} H_{N/2}^I & H_{N/2}^I \\ H_{N/2}^I T'_{N/2} & -H_{N/2}^I T'_{N/2} \end{pmatrix} \\
&= \frac{1}{\sqrt{2}} (I_1 \otimes H_{N/2}^I) \begin{pmatrix} I_{N/2} & I_{N/2} \\ T'_{N/2} & -T'_{N/2} \end{pmatrix} \\
(5.3) \quad &= \frac{1}{\sqrt{2}} (I_1 \otimes H_{N/2}^I) (I_{N/2} \oplus T'_{N/2}) \begin{pmatrix} I_{N/2} & I_{N/2} \\ I_{N/2} & -I_{N/2} \end{pmatrix} \\
&= (I_1 \otimes H_{N/2}^I) (I_{N/2} \oplus T'_{N/2}) (H_2^I \otimes I_{N/2}).
\end{aligned}$$

The latter concludes the proof.  $\square$

By transposing (5.2), we obtain an alternative orthogonal factorization

$$(5.4) \quad H_N^I = (H_2^I \otimes I_{N/2}) (I_{N/2} \oplus T'_{N/2}) (I_2 \otimes H_{N/2}^I) P_N.$$

EXAMPLE 5.4. For  $N = 4$  we observe that  $P_4^T = P_4$  and therefore, by Lemma 5.3, we have

$$2 H_4^I = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} = P_4 (I_2 \otimes \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}) \left( \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes I_2 \right).$$

For  $N = 8$ , Lemma 5.3 implies that

$$\sqrt{2} H_8^I = P_8^T (I_2 \otimes H_4^I) (I_4 \oplus T'_4) \left( \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes I_4 \right)$$

with

$$T'_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{2}/2 & 0 & \sqrt{2}/2 \\ 0 & 0 & 1 & 0 \\ 0 & \sqrt{2}/2 & 0 & -\sqrt{2}/2 \end{pmatrix}.$$

By means of the above factorization of  $2 H_4^I$ , we get the explicit factorization of  $\sqrt{8} H_8^I$  as

$$B_8 (I_4 \otimes \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}) (I_2 \otimes \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes I_2) (I_4 \oplus T'_4) \left( \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes I_4 \right).$$

ALGORITHM 5.5. (cas-I( $\mathbf{x}, N$ ))

<p>Input : <math>N = 2^t</math> (<math>t \geq 1</math>), <math>\mathbf{x} \in \mathbf{R}^N</math>.</p> <p>0. If <math>t \geq 3</math>, then precompute the nontrivial entries of <math>T'_n</math>, <math>n = 2^s</math>, (<math>s = 2, \dots, t-1</math>), i.e., <math>\cos(\pi j/N)</math>, (<math>j = 1, \dots, N/2-1</math>).</p> <p>1. If <math>t = 1</math>, then</p> $\mathbf{y} := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \mathbf{x}.$ <p>2. If <math>t = 2</math>, then</p> $(u_j)_{j=0}^3 := \left( \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes I_2 \right) \mathbf{x},$ $\mathbf{v}' := \text{cas} - \text{I}((u_j)_{j=0}^1, 2),$ $\mathbf{v}'' := \text{cas} - \text{I}((u_j)_{j=2}^3, 2),$ $\mathbf{y} := P_4 \begin{pmatrix} \mathbf{v}' \\ \mathbf{v}'' \end{pmatrix}.$ <p>3. If <math>t \geq 3</math>, then</p> $(u_j)_{j=0}^{N-1} := \left( \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes I_{N/2} \right) \mathbf{x},$ $(u'_j)_{j=N/2}^{N-1} := T'_{N/2} (u_j)_{j=N/2}^{N-1},$ $\mathbf{v}' := \text{cas} - \text{I}((u_j)_{j=0}^{N/2-1}, N/2),$ $\mathbf{v}'' := \text{cas} - \text{I}((u'_j)_{j=N/2}^{N-1}, N/2),$ $\mathbf{y} := P_N^T \begin{pmatrix} \mathbf{v}' \\ \mathbf{v}'' \end{pmatrix}.$
<p>Output : <math>\mathbf{y} = \sqrt{N} H_N^1 \mathbf{x}</math>.</p>

Note that  $B_8 = P_8^T (I_2 \otimes P_4)$  coincides with the classical *bit reversal matrix* of order 8 (see [39], pp. 36–43). If we compute  $\sqrt{8} H_8^1 \mathbf{x}$  for an arbitrary  $\mathbf{x} \in \mathbf{R}^8$ , then the algorithm based on the above factorization requires only 26 additions and 4 multiplications. Further we observe that only 12 butterfly operations and one rotation–reflection are used.

Let  $N = 2^t$ , ( $t \geq 1$ ) be a power of 2. Based on Lemma 5.3, we introduce a recursive procedure which is named cas-I( $\mathbf{x}, N$ ). The other algorithms called cas-II( $\mathbf{x}, N$ ), cas-III( $\mathbf{x}, N$ ), and cas-IV( $\mathbf{x}, N$ ) are variations of cas-I( $\mathbf{x}, N$ ) via Lemma 5.1. In order to reduce the number of multiplications, we move the factor  $1/\sqrt{2}$ , which appears in each application of Lemma 5.3 in the term  $H_2^1$  (see (5.3)), to the end of the calculation: in such a way, with given  $\mathbf{x} \in \mathbf{R}^N$ , we compute  $\mathbf{y} := \sqrt{N} H_N^X \mathbf{x}$  for  $X \in \{\text{I, II, III, IV}\}$ . By a final scaling defined as  $\mathbf{z} := \frac{1}{\sqrt{N}} \mathbf{y}$ , we obtain the transformed vector  $\mathbf{z} = H_N^X \mathbf{x}$ .

Applying Lemma 5.1, we devise the following algorithms for computing  $\mathbf{y} := \sqrt{N} H_N^X \mathbf{x}$  ( $X \in \{\text{II}, \text{III}, \text{IV}\}$ ) with arbitrary  $\mathbf{x} \in \mathbf{R}^N$ .

ALGORITHM 5.6. (cas-II( $\mathbf{x}, N$ ))

Input : $N = 2^t$ ( $t \geq 1$ ), $\mathbf{x} \in \mathbf{R}^N$ .
<p>0. If <math>t \geq 2</math>, then precompute the nontrivial entries of <math>T'_n</math>, <math>n = 2^s</math>, (<math>s = 2, \dots, t</math>), i.e., <math>\cos(\pi j/N)</math>, (<math>j = 1, \dots, N/2 - 1</math>).</p> <p>1. If <math>t = 1</math>, then</p> $\mathbf{y} := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \mathbf{x}.$ <p>2. If <math>t \geq 2</math>, then</p> $\mathbf{u} := \text{cas} - \text{I}(\mathbf{x}, N),$ $\mathbf{y} := T'_N \mathbf{u}.$
Output : $\mathbf{y} = \sqrt{N} H_N^{\text{II}} \mathbf{x}$ .

ALGORITHM 5.7. (cas-III( $\mathbf{x}, N$ ))

Input : $N = 2^t$ ( $t \geq 1$ ), $\mathbf{x} \in \mathbf{R}^N$ .
<p>0. If <math>t \geq 2</math>, then precompute the nontrivial entries of <math>T'_n</math>, <math>n = 2^s</math>, (<math>s = 2, \dots, t</math>), i.e., <math>\cos(\pi j/N)</math>, (<math>j = 1, \dots, N/2 - 1</math>).</p> <p>1. If <math>t = 1</math>, then</p> $\mathbf{y} := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \mathbf{x}.$ <p>2. If <math>t \geq 2</math>, then</p> $\mathbf{u} := T'_N \mathbf{x},$ $\mathbf{y} := \text{cas} - \text{I}(\mathbf{u}, N).$
Output : $\mathbf{y} = \sqrt{N} H_N^{\text{III}} \mathbf{x}$ .

ALGORITHM 5.8. (cas-IV( $\mathbf{x}$ ,  $N$ ))

Input : $N = 2^t$ ( $t \geq 1$ ), $\mathbf{x} \in \mathbf{R}^N$ .
0. If $t \geq 2$ , then precompute the nontrivial entries of $T'_n$ , $n = 2^s$ , ( $s = 2, \dots, t$ ), and $T_N$ , i.e., $\cos(\pi j/N)$ , ( $j = 1, \dots, N/2 - 1$ ) and $\cos((2j + 1)\pi/N)$ , ( $j = 1, \dots, N/2 - 1$ ). 1. If $t = 1$ , then $\mathbf{y} := \sqrt{2} \mathbf{x}$ . 2. If $t \geq 2$ , then $\mathbf{u} := T'_N \mathbf{x}$ , $\mathbf{v} := \text{cas} - \text{I}(\mathbf{u}, N)$ , $\mathbf{y} := T_N \mathbf{v}$ .
Output : $\mathbf{y} = \sqrt{N} H_N^{\text{IV}} \mathbf{x}$ .

Algorithms 5.5 – 5.8 are written in a short recursive form: this means that step 0 will be performed only once, at first (i.e. non recursive) call of cas- $X$ . In such algorithms we only use butterfly operations, rotation–reflections, and permutations. The number of arithmetic operations required for carrying out a computation is called *arithmetic cost*. Note that multiplications by  $2^k$  for some integer  $k$  and permutations are not counted. The aim of the rest of the section is the determination of the arithmetic costs of these fast DHT- $X(N)$  with  $X \in \{\text{I}, \text{II}, \text{III}, \text{IV}\}$ . For an arbitrary real matrix  $A_N$  of order  $N$ , let  $\alpha(A_N)$  and  $\mu(A_N)$  denote the number of additions and multiplications for computing  $A_N \mathbf{x}$  with any  $\mathbf{x} \in \mathbf{R}^N$ , respectively. Analogously, the number of additions and multiplications for computing of a fast DHT- $X(N)$  algorithm is denoted by  $\alpha(\text{DHT} - X, N)$  and  $\mu(\text{DHT} - X, N)$ , respectively.

Here we determine the arithmetic cost for computing  $\mathbf{y} = \sqrt{N} H_N^X \mathbf{x}$  with an arbitrary input vector  $\mathbf{x} \in \mathbf{R}^N$ .

**THEOREM 5.9.** *Let  $N = 2^t$  ( $t \geq 2$ ) be given. The arithmetic cost of the fast DHT-I( $N$ ) implemented by Algorithm 5.5 is described as follows*

$$\begin{aligned} \alpha(\text{DHT} - \text{I}, N) &= \frac{3}{2} N \log_2(N) - \frac{3}{2} N + 2, \\ \mu(\text{DHT} - \text{I}, N) &= N \log_2(N) - 3N + 4. \end{aligned}$$

*Proof.* For  $N = 2^t$  ( $t \geq 3$ ), by Algorithm 5.5 (or equivalently, by Lemma 5.3), we obtain the recursive relation

$$\begin{aligned} \alpha(\text{DHT} - \text{I}, N) &= 2 \alpha(\text{DHT} - \text{I}, N/2) + \alpha(T'_{N/2}) + N, \\ \mu(\text{DHT} - \text{I}, N) &= \mu(\text{DHT} - \text{I}, N/2) + \mu(T'_{N/2}), \end{aligned}$$

where, by definition of the twiddle matrix  $T'_N$  in Lemma 5.1, we have  $\alpha(T'_{N/2}) = N/2 - 2$  and  $\mu(T'_{N/2}) = 2(N/2 - 2) = N - 4$ . Thus we get the non-homogeneous linear (in the variable  $t$ ) difference equations of order 1

$$\begin{aligned}\alpha(\text{DHT} - \text{I}, 2^t) &= 2\alpha(\text{DHT} - \text{I}, 2^{t-1}) + 3 \times 2^{t-1} - 2, \\ \mu(\text{DHT} - \text{I}, 2^t) &= \mu(\text{DHT} - \text{I}, 2^{t-1}) + 2^t - 4.\end{aligned}$$

Solving the former difference equations under the initial conditions (see Example 5.4)

$$\alpha(\text{DHT} - \text{I}, 4) = 8, \quad \mu(\text{DHT} - \text{I}, 4) = 0,$$

we deduce that

$$\begin{aligned}\alpha(\text{DHT} - \text{I}, 2^t) &= 3t2^{t-1} - 3 \times 2^{t-1} + 2, \\ \mu(\text{DHT} - \text{I}, 2^t) &= t2^t - 3 \times 2^t + 4\end{aligned}$$

and this completes the proof.  $\square$

**COROLLARY 5.10.** *Let  $N = 2^t$  ( $t \geq 2$ ) be given. The arithmetic costs of the fast DHT- $X(N)$ ,  $X \in \{\text{II}, \text{III}, \text{IV}\}$ , implemented by Algorithms 5.6, 5.7, and 5.8 respectively, are described as follows*

$$\begin{aligned}\alpha(\text{DHT} - \text{II}, N) &= \alpha(\text{DHT} - \text{III}, N) = \frac{3}{2}N \log_2(N) - \frac{1}{2}N, \\ \mu(\text{DHT} - \text{II}, N) &= \mu(\text{DHT} - \text{III}, N) = N \log_2(N) - N, \\ \alpha(\text{DHT} - \text{IV}, N) &= \frac{3}{2}N \log_2(N) + \frac{1}{2}N, \quad \mu(\text{DHT} - \text{IV}, N) = N \log_2(N) + N.\end{aligned}$$

*Proof.* We apply Theorem 5.9 and we use the arithmetic costs of products with the twiddle matrices  $T'_N$  and  $T_N$  which, by their definition in Lemma 5.1, are given by

$$\alpha(T'_N) = N - 2, \quad \mu(T'_N) = 2N - 4, \quad \alpha(T_N) = N, \quad \mu(T_N) = 2N.$$

This completes the proof.  $\square$

**REMARK 5.11.** In [5], the authors provide further (substantially equivalent) factorizations of  $H_N^{\text{II}}$ ,  $H_N^{\text{III}}$ , and  $H_N^{\text{IV}}$  for  $N \geq 4$ . For example

$$H_N^{\text{II}} = P_N^T (I_2 \otimes H_{N/2}^{\text{II}}) (I_{N/2} \oplus T_{N/2}) (H_2^{\text{II}} \otimes I_{N/2}).$$

By making recourse to the factorizations in the case of  $N = 2^t$  ( $t \geq 2$ ), we can obtain corresponding fast DHT- $X(N)$  algorithms with  $X \in \{\text{II}, \text{III}, \text{IV}\}$  whose arithmetic costs are similar to those reported in Corollary 5.10.

A fast DHT algorithm is understood in the best way by interpreting it as the application of an explicit factorization of the corresponding Hartley matrix. We now

present a factorization of  $H_N^I$  with  $N = 2^t$  ( $t \geq 3$ ) into a product of sparse, orthogonal matrices. The considered factorization of  $H_N^I$  and Lemma 5.1 directly lead to direct DHT- $X(N)$  algorithms with  $X \in \{I, II, III, IV\}$ , which may be preferred on special platforms. Note that these direct formulations coincide with the recursive ones since they just arise from the explicit resolution of the recursions.

Let us consider  $H_N^I$  for  $N = 2^t$  ( $t \geq 3$ ). Further let  $N_s := 2^{t-s}$  ( $s = 0, \dots, t-1$ ). In the first factorization step, by Lemma 5.3, the matrix  $\sqrt{2} H_N^I$  is split in  $I_2 \otimes H_{N_1}^I$  by

$$\sqrt{2} H_N^I = P_N^T (I_2 \otimes H_{N_1}^I) (T_N^{(1)} A_N^{(1)})$$

with

$$A_N^{(1)} := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes I_{N_1}, \quad T_N^{(1)} := I_{N_1} \oplus T_{N_1}'.$$

In the second step we apply Lemma 5.3 to  $\sqrt{2} H_{N_1}^I$  and consequently we have

$$2 H_N^I = P_N^T (I_2 \otimes P_{N_1}^T) (I_4 \otimes H_{N_2}^I) (T_N^{(2)} A_N^{(2)}) (T_N^{(1)} A_N^{(1)})$$

with

$$A_N^{(2)} := I_2 \otimes A_{N_1}^{(1)}, \quad T_N^{(2)} := I_2 \otimes T_{N_1}^{(1)}.$$

In the case  $N_2 > 2$  we continue the procedure. Finally, after  $t-1$  steps, we obtain the factorization

$$(5.5) \quad 2^{t/2} H_N^I = B_N A_N^{(t)} A_N^{(t-1)} (T_N^{(t-2)} A_N^{(t-2)}) \dots (T_N^{(1)} A_N^{(1)}),$$

where

$$B_N := P_N^T (I_2 \otimes P_{N_1}^T) \dots (I_{N_{t-s}} \otimes P_{N_s}^T) \dots (I_{N_2} \otimes P_4^T)$$

is the *bit reversal matrix of order  $N$*  (see [39], pp. 36–43),

$$A_N^{(s)} := I_2 \otimes A_{N_1}^{(s-1)} = I_{N_{t+1-s}} \otimes A_{N_{s-1}}^{(1)} \quad (s = 2, \dots, t)$$

are modified addition matrices, and

$$T_N^{(s)} := I_2 \otimes T_{N_1}^{(1)} = I_{N_{t+1-s}} \otimes T_{N_{s-1}}^{(1)} \quad (s = 2, \dots, t-2)$$

are modified twiddle matrices. Note that  $(1/\sqrt{2}) A_N^{(s)}$  ( $s = 1, \dots, t$ ) and  $T_N^{(s)}$  ( $s = 1, \dots, t-2$ ) are orthogonal.  $A_N^{(s)}$  is called *almost orthogonal*, since it is a scaled orthogonal matrix. Furthermore, these matrices are sparse in the special sense that every row and column contains at most 2 nonzero entries. We summarize the previous analysis in the following theorem, whose proof follows directly from (5.5) and Lemma 5.1.

THEOREM 5.12. *Let  $N = 2^t$  ( $t \geq 3$ ) be given. Then the matrices  $H_N^X$  with  $X \in \{\text{I, II, III, IV}\}$  can be factorized into products of sparse, almost orthogonal matrices*

$$(5.6) \quad \begin{aligned} \sqrt{N} H_N^{\text{I}} &= B_N A_N^{(t)} A_N^{(t-1)} (T_N^{(t-2)} A_N^{(t-2)}) \cdots (T_N^{(1)} A_N^{(1)}), \\ \sqrt{N} H_N^{\text{II}} &= T_N' B_N A_N^{(t)} A_N^{(t-1)} (T_N^{(t-2)} A_N^{(t-2)}) \cdots (T_N^{(1)} A_N^{(1)}), \\ \sqrt{N} H_N^{\text{III}} &= B_N A_N^{(t)} A_N^{(t-1)} (T_N^{(t-2)} A_N^{(t-2)}) \cdots (T_N^{(1)} A_N^{(1)}) T_N', \\ \sqrt{N} H_N^{\text{IV}} &= T_N B_N A_N^{(t)} A_N^{(t-1)} (T_N^{(t-2)} A_N^{(t-2)}) \cdots (T_N^{(1)} A_N^{(1)}) T_N'. \end{aligned}$$

**6. Numerical stability of fast DHT algorithms.** In the following we use Wilkinson's standard model for the binary floating point arithmetic for real numbers (see [21], p. 44). If  $x \in \mathbf{R}$  is represented by the floating point number  $\hat{x} = \text{fl}(x)$ , then  $\text{fl}(x) = x(1 + \delta)$  with  $|\delta| \leq u$ , where  $u$  denotes the *unit roundoff* or *machine precision* as long as we disregard underflow and overflow. For arbitrary floating point numbers  $x_0, x_1$  and any arithmetical operation  $\odot \in \{+, -, \times, /\}$ , the exact value  $y = x_0 \odot x_1$  and the computed value  $\hat{y} = \text{fl}(x_0 \odot x_1)$  are related by

$$(6.1) \quad \text{fl}(x_0 \odot x_1) = (x_0 \odot x_1)(1 + \delta^\odot) \quad (|\delta^\odot| \leq u).$$

In the IEEE single precision arithmetic (24 bits for the mantissa including 1 sign bit, 8 bits for the exponent), we have  $u = 2^{-24} \approx 5.96 \cdot 10^{-8}$ . Concerning the double precision arithmetic (53 bits for the mantissa including 1 sign bit, 11 bits for the exponent), we find  $u = 2^{-53} \approx 1.11 \cdot 10^{-16}$  (see [21], p. 45).

The total roundoff error in the result of an algorithm is composed of a number of such errors when the input data are represented exactly. Otherwise we have also to add the inherent error: in our case, the orthogonality of the Hartley transforms guarantees that the inherent error is modest even for large length  $N$  (well conditioned problem) and then we can restrict our attention to the errors described in (6.1). To make clear the origin of the relative errors  $\delta_k^\odot$ , we use a superscript for specifying the operation  $\odot$  and a subscript for denoting the operation step  $k$ . In this section we show that under very mild assumptions, our fast DHT algorithms possess a remarkable numerical stability and this is of interest since the problem is well conditioned.

First we consider Algorithm 5.5 for DHT-I( $N$ ) with  $N = 2^t$  ( $t \geq 3$ ) which is equivalent to the factorization formula (5.6) of  $\sqrt{N} H_N^{\text{I}}$ . Our considerations are now based on the DHT-I( $N$ ) procedure in its direct form outlined below:

For  $s = 1$  to  $t - 2$  compute

$$(6.2) \quad \mathbf{x}^{(2s-1)} := A_N^{(s)} \mathbf{x}^{(2s-2)}, \quad \mathbf{x}^{(2s)} := T_N^{(s)} \mathbf{x}^{(2s-1)} \quad (\mathbf{x}^{(0)} := \mathbf{x})$$

and

$$(6.3) \quad \begin{aligned} \mathbf{x}^{(2t-3)} &:= A_N^{(t-1)} \mathbf{x}^{(2t-4)}, \quad \mathbf{x}^{(2t-2)} := A_N^{(t)} \mathbf{x}^{(2t-3)}, \\ \mathbf{x}^{(2t-1)} &:= B_N \mathbf{x}^{(2t-2)}. \end{aligned}$$

Then  $\mathbf{y} = \sqrt{N} H_N^I \mathbf{x} = \mathbf{x}^{(2t-1)}$  is the result of our fast DHT-I( $N$ ) Algorithm 5.5. By final scaling we obtain the wanted vector

$$\mathbf{z} = H_N^I \mathbf{x} = \frac{1}{\sqrt{N}} \mathbf{x}^{(2t-1)}.$$

The roundoff errors of the considered Algorithm 5.5 are caused by the modified addition matrices  $A_N^{(s)}$  ( $s = 1, \dots, t$ ), the modified twiddle matrices  $T_N^{(s)}$  ( $s = 1, \dots, t-2$ ), and the final scaling. The matrices  $A_N^{(s)}$  and  $T_N^{(s)}$  possess a simple structure. After suitable permutations, every of such matrices is block diagonal with block order at most 2. All blocks of order 1 coincide with 1, which correspond to identities and therefore do not produce roundoff errors. Every block of order 2 is either a butterfly matrix

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

or a rotation-reflection matrix

$$\begin{pmatrix} a_0 & a_1 \\ a_1 & -a_0 \end{pmatrix}$$

with  $a_0 = \cos(2^{-s}\pi j)$ ,  $a_1 = \sin(2^{-s}\pi j)$ , ( $s = 2, \dots, t-1$ ;  $j = 1, \dots, 2^{s-1} - 1$ ). The following result [28] is of interest for our analysis.

LEMMA 6.1. *Let  $x_0, x_1$  be arbitrary real floating point numbers.*

1. *For the butterfly operation,  $y_0 := x_0 + x_1$ ,  $y_1 := x_0 - x_1$  with  $\hat{y}_0 := \text{fl}(x_0 + x_1)$ ,  $\hat{y}_1 := \text{fl}(x_0 - x_1)$ , the roundoff error can be estimated as*

$$(\hat{y}_0 - y_0)^2 + (\hat{y}_1 - y_1)^2 \leq 2u^2(x_0^2 + x_1^2).$$

2. *Let  $s \geq 2$  be a given integer. If  $a_0 = \cos(2^{-s}\pi j)$  and  $a_1 = \sin(2^{-s}\pi j)$ , ( $j = 1, \dots, 2^{s-1} - 1$ ) are precomputed by  $\hat{a}_k = a_k + \Delta a_k$  with  $|\Delta a_k| \leq cu$  ( $k = 0, 1$ ), then for the rotation-reflection  $y_0 := a_0 x_0 + a_1 x_1$ ,  $y_1 := a_1 x_0 - a_0 x_1$  with  $\hat{y}_0 := \text{fl}(a_0 x_0 + a_1 x_1)$ ,  $\hat{y}_1 := \text{fl}(a_1 x_0 - a_0 x_1)$ , the roundoff error can be estimated as*

$$(\hat{y}_0 - y_0)^2 + (\hat{y}_1 - y_1)^2 \leq \left(\frac{4}{3}\sqrt{3} + c\sqrt{2} + O(u)\right)^2 u^2 (x_0^2 + x_1^2).$$

We consider arbitrary input vectors  $\mathbf{x} \in \mathbf{R}^N$ , where all components of  $\mathbf{x}$  are floating point numbers. In this way we neglect the inherent error and we essentially consider only the algorithmic one. Let  $\mathbf{y} = \sqrt{N} H_N^X \mathbf{x} \in \mathbf{R}^N$  with  $X \in \{\text{I, II, III, IV}\}$  such that  $\mathbf{z} = \frac{1}{\sqrt{N}} \mathbf{y}$  is the exact transformed vector. Further let  $\hat{\mathbf{y}}$  the output vector computed by using Algorithms 5.5, 5.6, 5.7, and 5.8, respectively, using a binary floating point arithmetic with unit roundoff  $u$ . Finally, let  $\hat{\mathbf{z}} := \text{fl}\left(\frac{1}{\sqrt{N}} \mathbf{y}\right)$ . Since  $H_N^X$  is invertible,  $\hat{\mathbf{z}}$  can be represented in the form  $\hat{\mathbf{z}} = H_N^X (\mathbf{x} + \Delta \mathbf{x})$  with  $\Delta \mathbf{x} \in \mathbf{R}^N$ . An

algorithm for computing  $H_N^X \mathbf{x}$  is called *normwise backward stable* (see [21], p. 142), if there exists a positive constant  $k_N$  with  $k_N u \ll 1$  such that

$$(6.4) \quad \|\Delta \mathbf{x}\|_2 \leq (k_N u + O(u^2)) \|\mathbf{x}\|_2$$

for all input vectors  $\mathbf{x}$ . Thus the constant  $k_N$  measures the numerical stability of the algorithm. Since  $H_N^X$  is orthogonal, we conclude that  $\|\Delta \mathbf{x}\|_2 = \|H_N^X(\Delta \mathbf{x})\|_2 = \|\hat{\mathbf{z}} - \mathbf{z}\|_2$  and  $\|\mathbf{x}\|_2 = \|H_N^X \mathbf{x}\|_2 = \|\mathbf{z}\|_2$ . Hence, if (6.4) is satisfied, then we have *normwise forward stability* as well, i.e.

$$\|\hat{\mathbf{z}} - \mathbf{z}\|_2 \leq (k_N u + O(u^2)) \|\mathbf{z}\|_2.$$

Let  $N = 2^t$  ( $t \geq 3$ ) be given. We now look closer at the computational steps in our recursive DHT-I( $N$ ) Algorithm 5.5 which is equivalent to its direct version (6.2) – (6.3). For every  $s = 2, \dots, t-1$ , all values

$$(6.5) \quad \cos(2^{-s}\pi j) \quad (j = 1, \dots, 2^{s-1} - 1)$$

needed in the twiddle matrix  $T'_{2^s}$  are precomputed. Note that

$$\sin(2^{-s}\pi j) = \cos(2^{-s}\pi(2^{s-1} - j)) \quad (j = 1, \dots, 2^{s-1} - 1).$$

If cosine is internally computed with high precision and all values (6.5) afterwards are rounded to the next machine precision, then we obtain a very accurate approximation of the quantities (6.5) with an error constant  $c = 1/2$  (see Lemma 6.1, item 2). In the following, we use the matrix  $\hat{T}_N^{(s)}$  ( $s = 1, \dots, t-2$ ) with the precomputed entries (6.5) instead of the exact modified twiddle matrices  $T_N^{(s)}$ . Set  $\hat{\mathbf{x}}^{(0)} = \mathbf{x}^{(0)} = \mathbf{x}$ . Then the vectors computed in the direct DHT-I( $N$ ) algorithm (6.2) – (6.3), denoted by  $\hat{\mathbf{x}}^{(r)}$  ( $r = 1, \dots, 2t-1$ ), satisfy the following relationships:

$$\begin{aligned} \hat{\mathbf{x}}^{(2s-1)} &:= \text{fl} \left( A_N^{(s)} \hat{\mathbf{x}}^{(2s-2)} \right) \quad (s = 1, \dots, t-2), \\ \hat{\mathbf{x}}^{(2s)} &:= \text{fl} \left( \hat{T}_N^{(s)} \hat{\mathbf{x}}^{(2s-1)} \right) \quad (s = 1, \dots, t-2), \\ \hat{\mathbf{x}}^{(2t-3)} &:= \text{fl} \left( A_N^{(t-1)} \hat{\mathbf{x}}^{(2t-4)} \right), \\ \hat{\mathbf{x}}^{(2t-2)} &:= \text{fl} \left( A_N^{(t)} \hat{\mathbf{x}}^{(2t-3)} \right), \\ \hat{\mathbf{x}}^{(2t-1)} &:= B_N \hat{\mathbf{x}}^{(2t-2)}. \end{aligned}$$

Further we introduce (implicitly) the error vectors  $\mathbf{e}^{(r)} \in \mathbf{R}^N$  ( $r = 1, \dots, 2t-1$ ) by

$$(6.6) \quad \begin{aligned} \hat{\mathbf{x}}^{(2s-1)} &= A_N^{(s)} \hat{\mathbf{x}}^{(2s-2)} + \mathbf{e}^{(2s-1)} \quad (s = 1, \dots, t-2), \\ \hat{\mathbf{x}}^{(2s)} &= T_N^{(s)} \hat{\mathbf{x}}^{(2s-1)} + \mathbf{e}^{(2s)} \quad (s = 1, \dots, t-2), \\ \hat{\mathbf{x}}^{(2t-3)} &= A_N^{(t-1)} \hat{\mathbf{x}}^{(2t-4)} + \mathbf{e}^{(2t-3)}, \\ \hat{\mathbf{x}}^{(2t-2)} &= A_N^{(t)} \hat{\mathbf{x}}^{(2t-3)} + \mathbf{e}^{(2t-2)}, \\ \hat{\mathbf{x}}^{(2t-1)} &= B_N \hat{\mathbf{x}}^{(2t-2)} + \mathbf{e}^{(2t-1)}. \end{aligned}$$

Note that  $\mathbf{e}^{(r)}$  ( $r = 1, \dots, 2t-1$ ) describes the precomputation error and the roundoff error at the  $r$ -th step in DHT-I( $N$ ) procedure. Since  $B_N$  is the bit reversal matrix, we have  $\mathbf{e}^{(2t-1)} = \mathbf{o}$ . The matrix vector products  $A_N^{(s)} \hat{\mathbf{x}}^{(2s-2)}$  ( $s = 1, \dots, t-1$ ), and  $A_N^{(t)} \hat{\mathbf{x}}^{(2t-3)}$  involve only butterfly operations and consequently by Lemma 6.1, item 1, we have

$$(6.7) \quad \begin{aligned} \|\mathbf{e}^{(2s-1)}\|_2 &\leq \sqrt{2} u \|\hat{\mathbf{x}}^{(2s-2)}\|_2 \quad (s = 1, \dots, t-1), \\ \|\mathbf{e}^{(2t-2)}\|_2 &\leq \sqrt{2} u \|\hat{\mathbf{x}}^{(2t-3)}\|_2. \end{aligned}$$

Every matrix vector product  $T_N^{(s)} \hat{\mathbf{x}}^{(2s-1)}$  ( $s = 1, \dots, t-2$ ) consists of identities and rotation-reflections and then by Lemma 6.1, item 2, we deduce the estimate

$$(6.8) \quad \|\mathbf{e}^{(2s)}\|_2 \leq \left(\frac{4}{3}\sqrt{3} + \frac{1}{2}\sqrt{2} + O(u)\right) u \|\hat{\mathbf{x}}^{(2s-1)}\|_2, \quad (s = 1, \dots, t-2).$$

Finally, we scale the result of our fast DHT-I( $N$ ) algorithm by the relation  $\mathbf{z} = 2^{-t/2} \mathbf{x}^{(2t-1)}$ . Let  $\hat{\mathbf{z}} = \text{fl}(2^{-t/2} \hat{\mathbf{x}}^{(2t-1)})$ . For even  $t \geq 3$ , the latter scaling by a power of 2 does not lead to any additional roundoff error and hence

$$\|\hat{\mathbf{z}} - \mathbf{z}\|_2 = 2^{-t/2} \|\hat{\mathbf{x}}^{(2t-1)} - \mathbf{x}^{(2t-1)}\|_2 = 2^{-t/2} \|\hat{\mathbf{x}}^{(2t-2)} - \mathbf{x}^{(2t-2)}\|_2.$$

For odd  $t \geq 3$ , we precompute  $\text{fl}(2^{-t/2})$  as  $2^{-(t+1)/2} \text{fl}(\sqrt{2})$  with  $|\text{fl}(\sqrt{2}) - \sqrt{2}| \leq u$ . Then by (6.1), for every  $j = 0, \dots, N-1$ , we obtain

$$\hat{z}_j = 2^{-(t+1)/2} \text{fl}(\sqrt{2}) \hat{x}^{(2t-1)}(1 + \delta_j^\times)$$

with  $|\delta_j^\times| \leq u$  and the latter implies that

$$\|\hat{\mathbf{z}} - 2^{-t/2} \mathbf{x}^{(2t-1)}\|_2 \leq 2^{-t/2} \|\mathbf{x}^{(2t-1)}\|_2 u \left(1 + \frac{\sqrt{2}}{2} + O(u)\right).$$

Hence by triangle inequality it follows that

$$(6.9) \quad \begin{aligned} \|\hat{\mathbf{z}} - \mathbf{z}\|_2 &\leq \frac{1}{\sqrt{N}} \|\mathbf{x}^{(2t-1)}\|_2 u \left(1 + \frac{\sqrt{2}}{2} + O(u)\right) + \frac{1}{\sqrt{N}} \|\hat{\mathbf{x}}^{(2t-1)} - \mathbf{x}^{(2t-1)}\|_2 \\ &= \frac{1}{\sqrt{N}} \|\mathbf{x}^{(2t-1)}\|_2 u \left(1 + \frac{\sqrt{2}}{2} + O(u)\right) + \frac{1}{\sqrt{N}} \|\hat{\mathbf{x}}^{(2t-1)} - \mathbf{x}^{(2t-1)}\|_2 \end{aligned}$$

which is true for even  $t \geq 3$  too.

Now we are ready for estimating the total roundoff error  $\|\hat{\mathbf{z}} - \mathbf{z}\|_2$  of the presented fast DHT-I( $N$ ) algorithm under the assumption that  $\sqrt{2}$  and the cosine values (6.5) are precomputed in the best possible way.

**THEOREM 6.2.** *Let  $N = 2^t$  ( $t \geq 3$ ) be given. Assume that  $\sqrt{2}$  and the values (6.5) are precomputed with an absolute error bound of  $u$  and  $u/2$ , respectively. Then the fast DHT-I( $N$ ) Algorithm 5.5 (with final scaling) is normwise backward stable with the constant*

$$k_N = \left(\frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2}\right) (\log_2(N) - 1) \approx 4.430721 (\log_2(N) - 1).$$

*Proof.* 1. First we estimate the roundoff error  $\|\hat{\mathbf{x}}^{(2t-1)} - \mathbf{x}^{(2t-1)}\|_2$ . Applying (6.6) and (6.2) – (6.3) repeatedly, we find that

$$(6.10) \quad \begin{aligned} \hat{\mathbf{x}}^{(2t-1)} &= \mathbf{x}^{(2t-1)} + B_N A_N^{(t)} A_N^{(t-1)} T_N^{(t-2)} \dots T_N^{(2)} A_N^{(2)} T_N^{(1)} \mathbf{e}^{(1)} \\ &\quad + B_N A_N^{(t)} A_N^{(t-1)} T_N^{(t-2)} \dots T_N^{(2)} A_N^{(2)} \mathbf{e}^{(2)} + \dots \\ &\quad + B_N A_N^{(t)} A_N^{(t-1)} T_N^{(t-2)} \mathbf{e}^{(2t-5)} + B_N A_N^{(t)} A_N^{(t-1)} \mathbf{e}^{(2t-4)} \\ &\quad + B_N A_N^{(t)} \mathbf{e}^{(2t-3)} + B_N \mathbf{e}^{(2t-2)}. \end{aligned}$$

The matrices  $\frac{1}{\sqrt{2}} A_N^{(s)}$ , ( $s = 1, \dots, t$ ) and  $T_N^{(s)}$ , ( $s = 1, \dots, t-2$ ) are orthogonal and as a consequence all matrices have the spectral norm equal to 1, i.e.,  $\|A_N^{(s)}\|_2 = \sqrt{2}$ , ( $s = 1, \dots, t$ ) and  $\|T_N^{(s)}\|_2 = 1$ , ( $s = 1, \dots, t-2$ ). Then by (6.6) we deduce the estimates

$$\begin{aligned} \|\hat{\mathbf{x}}^{(2s-1)}\|_2 &\leq \sqrt{2} \|\hat{\mathbf{x}}^{(2s-2)}\|_2 + \|\mathbf{e}^{(2s-1)}\|_2 \quad (s = 1, \dots, t-1), \\ \|\hat{\mathbf{x}}^{(2s)}\|_2 &\leq \|\hat{\mathbf{x}}^{(2s-1)}\|_2 + \|\mathbf{e}^{(2s)}\|_2 \quad (s = 1, \dots, t-2), \\ \|\hat{\mathbf{x}}^{(2t-1)}\|_2 &= \|\hat{\mathbf{x}}^{(2t-2)}\|_2 \leq \sqrt{2} \|\hat{\mathbf{x}}^{(2t-3)}\|_2 + \|\mathbf{e}^{(2t-2)}\|_2. \end{aligned}$$

Thus by (6.7) – (6.8) we infer that

$$\begin{aligned} \|\hat{\mathbf{x}}^{(2s-1)}\|_2 &\leq (\sqrt{2} + O(u)) \|\hat{\mathbf{x}}^{(2s-2)}\|_2 \quad (s = 1, \dots, t-1), \\ \|\hat{\mathbf{x}}^{(2s)}\|_2 &\leq (1 + O(u)) \|\hat{\mathbf{x}}^{(2s-1)}\|_2 \quad (s = 1, \dots, t-2), \\ \|\hat{\mathbf{x}}^{(2t-1)}\|_2 &= \|\hat{\mathbf{x}}^{(2t-2)}\|_2 \leq (\sqrt{2} + O(u)) \|\hat{\mathbf{x}}^{(2t-3)}\|_2. \end{aligned}$$

Since  $\hat{\mathbf{x}}^{(0)} = \mathbf{x}$ , we have

$$(6.11) \quad \begin{aligned} \|\hat{\mathbf{x}}^{(s)}\|_2 &\leq (2^{\lceil s/2 \rceil} + O(u)) \|\mathbf{x}\|_2 \quad (s = 1, \dots, 2t-3), \\ \|\hat{\mathbf{x}}^{(2t-1)}\|_2 &= \|\hat{\mathbf{x}}^{(2t-2)}\|_2 \leq (2^{t/2} + O(u)) \|\mathbf{x}\|_2. \end{aligned}$$

From (6.7) – (6.8) and (6.11), it follows that

$$(6.12) \quad \begin{aligned} \|\mathbf{e}^{(2s-1)}\|_2 &\leq (2^{(s+1)/2} + O(u)) u \|\mathbf{x}\|_2 \quad (s = 1, \dots, t-1), \\ \|\mathbf{e}^{(2s)}\|_2 &\leq 2^{s/2} \left( \frac{4}{3} \sqrt{3} + \frac{1}{2} \sqrt{2} + O(u) \right) u \|\mathbf{x}\|_2 \quad (s = 1, \dots, t-2), \\ \|\mathbf{e}^{(2t-2)}\|_2 &\leq \left( 2^{s/2} + O(u) \right) u \|\mathbf{x}\|_2. \end{aligned}$$

2. By (6.10) we directly find

$$\begin{aligned} \|\hat{\mathbf{x}}^{(2t-1)} - \mathbf{x}^{(2t-1)}\|_2 &= \|A_N^{(t)}\|_2 \|A_N^{(t-1)}\|_2 \dots \|A_N^{(2)}\|_2 \|\mathbf{e}^{(1)}\|_2 \\ &\quad + \|A_N^{(t)}\|_2 \|A_N^{(t-1)}\|_2 \dots \|A_N^{(2)}\|_2 \|\mathbf{e}^{(2)}\|_2 + \dots \\ &\quad + \|A_N^{(t)}\|_2 \|A_N^{(t-1)}\|_2 \|\mathbf{e}^{(2t-5)}\|_2 + \|A_N^{(t)}\|_2 \|A_N^{(t-1)}\|_2 \|\mathbf{e}^{(2t-4)}\|_2 \\ &\quad + \|A_N^{(t)}\|_2 \|\mathbf{e}^{(2t-3)}\|_2 + \|\mathbf{e}^{(2t-2)}\|_2. \end{aligned}$$

Hence (6.12) implies the conclusion

$$(6.13) \quad \|\hat{\mathbf{x}}^{(2t-1)} - \mathbf{x}^{(2t-1)}\|_2 \leq 2^{t/2} \left[ \left( \frac{4}{3} \sqrt{3} + \frac{3}{2} \sqrt{2} \right) (t-2) + \sqrt{2} + 1 + O(u) \right] u \|\mathbf{x}\|_2.$$

3. For the final scaling described by  $\mathbf{z} = 2^{-t/2}\mathbf{x}^{(2t-1)}$  we set  $\hat{\mathbf{z}} = \text{fl}(2^{-t/2}\hat{\mathbf{x}}^{(2t-1)})$ . By using (6.9) we obtain the final result

$$\|\hat{\mathbf{z}} - \mathbf{z}\|_2 \leq \left[ \left( \frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2} \right) (t-1) + O(u) \right] u \|\mathbf{x}\|_2$$

and the theorem is proven.  $\square$

**COROLLARY 6.3.** *Let  $N = 2^t$  ( $t \geq 3$ ) be given. Assume that  $\sqrt{2}$  and the values (6.5) are precomputed with an absolute error bound of  $u$  and  $u/2$ , respectively. Then the fast DHT-II( $N$ ) Algorithm 5.6 and the fast DHT-III( $N$ ) Algorithm 5.7 (with final scaling) are normwise backward stable with the constant*

$$k_N = \left( \frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2} \right) \log_2(N) - \sqrt{2} \approx 4.430721 \log_2(N) - 1.414214.$$

*Proof.* We apply mainly the new result (6.13) from the proof of Theorem 6.2. Taking into account the notations of Algorithm 5.6, we consider the exact vectors  $\mathbf{u} = \sqrt{N} H_N^I \mathbf{x}$ ,  $\mathbf{y} = T'_N \mathbf{u}$ ,  $\mathbf{z} = \frac{1}{\sqrt{N}} \mathbf{y} = H_N^{II} \mathbf{x}$ , and the corresponding computed vectors  $\hat{\mathbf{u}}, \hat{\mathbf{y}} := \text{fl}(T'_N \hat{\mathbf{u}})$ ,  $\hat{\mathbf{z}} := \text{fl}(\frac{1}{\sqrt{N}} \hat{\mathbf{y}})$ . We introduce the error vectors  $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}$  through the relations

$$\hat{\mathbf{u}} = \mathbf{u} + \mathbf{e}^{(1)}, \quad \hat{\mathbf{y}} = T'_N \hat{\mathbf{u}} + \mathbf{e}^{(2)}$$

such that

$$\hat{\mathbf{y}} = T'_N \mathbf{u} + T'_N \mathbf{e}^{(1)} + \mathbf{e}^{(2)} = \mathbf{y} + T'_N \mathbf{e}^{(1)} + \mathbf{e}^{(2)}.$$

Since  $T'_N$  is orthogonal, we get  $\|T'_N\|_2 = 1$  and hence by triangle inequality

$$(6.14) \quad \|\hat{\mathbf{y}} - \mathbf{y}\|_2 \leq \|\mathbf{e}^{(1)}\|_2 + \|\mathbf{e}^{(2)}\|_2.$$

By (6.13) we see that

$$(6.15) \quad \|\mathbf{e}^{(1)}\|_2 \leq \sqrt{N} \left[ \left( \frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2} \right) (t-2) + \sqrt{2} + 1 + O(u) \right] u \|\mathbf{x}\|_2.$$

By virtue of Lemma 6.1, item 2, it follows that

$$\|\mathbf{e}^{(2)}\|_2 \leq \left( \frac{4}{3}\sqrt{3} + \frac{1}{2}\sqrt{2} + O(u) \right) u \|\hat{\mathbf{u}}\|_2,$$

since, after suitable permutations,  $T'_N$  is the direct sum of one identity matrix and  $N/2 - 1$  rotation-reflection matrices of order 2. Note that by (6.15)

$$\|\hat{\mathbf{u}}\|_2 \leq \|\mathbf{u}\|_2 + \|\mathbf{e}^{(1)}\|_2 = \|\sqrt{N} H_N^I \mathbf{x}\|_2 + \|\mathbf{e}^{(1)}\|_2 \leq (\sqrt{N} + O(u)) \|\mathbf{x}\|_2$$

and consequently

$$(6.16) \quad \|\mathbf{e}^{(2)}\|_2 \leq \sqrt{N} \left( \frac{4}{3}\sqrt{3} + \frac{1}{2}\sqrt{2} + O(u) \right) u \|\mathbf{x}\|_2.$$

Then (6.14) – (6.16) imply that

$$\|\hat{\mathbf{y}} - \mathbf{y}\|_2 \leq \sqrt{N} \left[ \left( \frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2} \right) (t-1) + 1 + O(u) \right] u \|\mathbf{x}\|_2.$$

Note that by (6.16) we have

$$\begin{aligned} \|\hat{\mathbf{y}}\|_2 &= \|T'_N\|_2 \|\hat{\mathbf{u}}\|_2 + \|\mathbf{e}^{(2)}\|_2 \\ &\leq \|\hat{\mathbf{u}}\|_2 + \|\mathbf{e}^{(2)}\|_2 \\ &\leq (\sqrt{N} + O(u)) \|\mathbf{x}\|_2. \end{aligned}$$

By recalling (6.9), for the final scaling we infer that

$$\begin{aligned} \|\hat{\mathbf{z}} - \mathbf{z}\|_2 &\leq \frac{1}{\sqrt{N}} \|\hat{\mathbf{y}}\|_2 u \left( 1 + \frac{\sqrt{2}}{2} + O(u) \right) + \frac{1}{\sqrt{N}} \|\hat{\mathbf{y}} - \mathbf{y}\|_2 \\ &\leq \left[ \left( \frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2} \right) t - \sqrt{2} + O(u) \right] u \|\mathbf{x}\|_2. \end{aligned}$$

The proof for Algorithm 5.7 is totally similar and is omitted here.  $\square$

**COROLLARY 6.4.** *Let  $N = 2^t$  ( $t \geq 3$ ) be given. Assume that  $\sqrt{2}$  is precomputed with an absolute error bound of  $u$  and the values  $\cos(2^{-s}\pi j)$  ( $s = 1, \dots, t; j = 1, \dots, 2^{s-1} - 1$ ) and  $\cos(2^{-t-1}\pi(2j+1))$  ( $j = 0, \dots, 2^{t-1} - 1$ ) are precomputed with an absolute error bound of  $u/2$ . Then the fast DHT-IV( $N$ ) Algorithm 5.8 (with final scaling) is normwise backward stable with the constant*

$$k_N = \left( \frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2} \right) \log_2(N) + 1 \approx 4.430721 \log_2(N) + 1.$$

*Proof.* By using the notations of Algorithm 5.8, we consider the exact vectors  $\mathbf{u} = T'_N \mathbf{x}$ ,  $\mathbf{v} = \sqrt{N} H_N^I \mathbf{u}$ ,  $\mathbf{y} = T_N \mathbf{v}$ ,  $\mathbf{z} = \frac{1}{\sqrt{N}} \mathbf{y} = H_N^{IV} \mathbf{x}$ , and the corresponding computed vectors  $\hat{\mathbf{u}} := \text{fl}(\hat{T}'_N \hat{\mathbf{x}})$ ,  $\hat{\mathbf{v}}, \hat{\mathbf{y}} := \text{fl}(\hat{T}_N \hat{\mathbf{v}})$ ,  $\hat{\mathbf{z}} := \text{fl}(\frac{1}{\sqrt{N}} \hat{\mathbf{y}})$ . We introduce the error vectors  $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \mathbf{e}^{(3)}$  via the identities

$$\hat{\mathbf{u}} = T'_N \mathbf{x} + \mathbf{e}^{(1)}, \quad \hat{\mathbf{v}} = \sqrt{N} H_N^I \hat{\mathbf{u}} + \mathbf{e}^{(2)}, \quad \hat{\mathbf{y}} = T_N \hat{\mathbf{v}} + \mathbf{e}^{(3)}$$

and hence by Lemma 5.1

$$\begin{aligned} \hat{\mathbf{y}} &= \sqrt{N} T_N H_N^I T'_N \mathbf{x} + \sqrt{N} T_N H_N^I \mathbf{e}^{(1)} + T_N \mathbf{e}^{(2)} + \mathbf{e}^{(3)} \\ &= \mathbf{y} + \sqrt{N} T_N H_N^I \mathbf{e}^{(1)} + T_N \mathbf{e}^{(2)} + \mathbf{e}^{(3)}. \end{aligned}$$

Then, by triangle inequality and by orthogonality of  $H_N^I$  and  $T_N$ , we find

$$\begin{aligned} \|\hat{\mathbf{y}} - \mathbf{y}\|_2 &\leq \|\sqrt{N} T_N H_N^I \mathbf{e}^{(1)}\|_2 + \|T_N \mathbf{e}^{(2)}\|_2 + \|\mathbf{e}^{(3)}\|_2 \\ (6.17) \quad &\leq \sqrt{N} \|\mathbf{e}^{(1)}\|_2 + \|\mathbf{e}^{(2)}\|_2 + \|\mathbf{e}^{(3)}\|_2. \end{aligned}$$

By (6.13) we see that

$$\|\mathbf{e}^{(2)}\|_2 \leq \sqrt{N} \left[ \left( \frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2} \right) (t-2) + \sqrt{2} + 1 + O(u) \right] u \|\hat{\mathbf{u}}\|_2.$$

From Lemma 6.1, item 2, we deduce the following relations

$$(6.18) \quad \begin{aligned} \|\mathbf{e}^{(1)}\|_2 &\leq \left(\frac{4}{3}\sqrt{3} + \frac{1}{2}\sqrt{2} + O(u)\right) u \|\mathbf{x}\|_2, \\ \|\mathbf{e}^{(3)}\|_2 &\leq \left(\frac{4}{3}\sqrt{3} + \frac{1}{2}\sqrt{2} + O(u)\right) u \|\hat{\mathbf{v}}\|_2. \end{aligned}$$

Since

$$\|\hat{\mathbf{u}}\|_2 \leq \|\mathbf{x}\|_2 + \|\mathbf{e}^{(1)}\|_2 \leq (1 + O(u)) \|\mathbf{x}\|_2,$$

we infer

$$(6.19) \quad \|\mathbf{e}^{(2)}\|_2 \leq \sqrt{N} \left[ \left(\frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2}\right) (t - 2) + \sqrt{2} + 1 + O(u) \right] u \|\mathbf{x}\|_2$$

and then we have the estimate

$$\begin{aligned} \|\hat{\mathbf{v}}\|_2 &\leq \|\sqrt{N} H_N^I \hat{\mathbf{u}}\|_2 + \|\mathbf{e}^{(2)}\|_2 = \sqrt{N} \|\hat{\mathbf{u}}\|_2 + \|\mathbf{e}^{(2)}\|_2 \\ &\leq (\sqrt{N} + O(u)) \|\mathbf{x}\|_2. \end{aligned}$$

The latter implies that

$$(6.20) \quad \|\mathbf{e}^{(3)}\|_2 \leq \sqrt{N} \left(\frac{4}{3}\sqrt{3} + \frac{1}{2}\sqrt{2} + O(u)\right) u \|\mathbf{x}\|_2,$$

$$\|\hat{\mathbf{y}}\|_2 \leq \|T_N \hat{\mathbf{v}}\|_2 + \|\mathbf{e}^{(3)}\|_2 = \|\hat{\mathbf{v}}\|_2 + \|\mathbf{e}^{(3)}\|_2$$

$$(6.21) \quad \leq (\sqrt{N} + O(u)) \|\mathbf{x}\|_2.$$

Then (6.17) – (6.20) imply that

$$\|\hat{\mathbf{y}} - \mathbf{y}\|_2 \leq \sqrt{N} \left[ \left(\frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2}\right) t - \sqrt{2} + 1 + O(u) \right] u \|\mathbf{x}\|_2.$$

By recalling (6.9) and using (6.21), for the final scaling we obtain the estimate

$$\begin{aligned} \|\hat{\mathbf{z}} - \mathbf{z}\|_2 &\leq \frac{1}{\sqrt{N}} \|\hat{\mathbf{y}}\|_2 u \left(1 + \frac{\sqrt{2}}{2} + O(u)\right) + \frac{1}{\sqrt{N}} \|\hat{\mathbf{y}} - \mathbf{y}\|_2 \\ &\leq \left[ \left(\frac{4}{3}\sqrt{3} + \frac{3}{2}\sqrt{2}\right) t + 1 + O(u) \right] u \|\mathbf{x}\|_2. \end{aligned}$$

This completes the proof.  $\square$

**7. Optimal Frobenius approximation of Toeplitz matrices.** In this section we consider the approximation of Toeplitz structures by matrices in the four Hartley matrix algebras described in Section 4. Since all the matrices of a Hartley matrix algebra are inherently real symmetric, we restrict our attention to real symmetric Toeplitz matrices. More precisely, we consider symmetric Toeplitz matrices  $T_N(f)$  generated by a Lebesgue integrable even function  $f : (-\pi, \pi] \rightarrow \mathbf{R}$  in the sense that the entries of  $T_N(f)$  along the  $k$ -th diagonal are given by the  $k$ -th Fourier coefficient  $t_k$  of  $f$ :

$$(7.1) \quad [T_N(f)]_{i,j} = t_{|i-j|}, \quad t_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx = \frac{1}{\pi} \int_0^{\pi} f(x) \cos(kx) dx \quad (k \in \mathbf{Z}).$$

Then the function  $f$  is called *symbol* or *generating function* of  $T_N(f)$ .

Let  $U_N$  be given unitary matrices for all orders  $N \in \mathbf{N}$  and  $U_N^* := \overline{U_N}^T$  their complex conjugates, where the case  $U_N = H_N^X$  with  $X \in \{\text{I, II, III, IV}\}$  is of special interest. By  $\text{Diag}(U_N) := \{A \in \mathbf{C}^{N \times N} : U_N^* A U_N \text{ diagonal}\}$  we denote the matrix algebra related to  $U_N$ . Now we introduce the operator  $\mathcal{P}_{U_N} : \mathbf{C}^{N \times N} \rightarrow \text{Diag}(U_N)$  by  $\mathcal{P}_{U_N}(A) = \tilde{A}$  for any  $A \in \mathbf{C}^{N \times N}$ , where  $\tilde{A} \in \text{Diag}(U_N)$  minimizes the quadratic functional

$$\mathcal{F}_A(X) := \langle A - X, A - X \rangle = \|A - X\|_F^2 \quad (X \in \text{Diag}(U_N))$$

in the Frobenius norm (associated to the Frobenius inner product (4.1)). In this way,  $\mathcal{P}_{U_N}(T_N(f))$  is that matrix where the functional  $\mathcal{F}_{T_N(f)}$  attains its minimum value. Then  $\mathcal{P}_{U_N}(T_N(f))$  is called the *optimal preconditioner* of the Toeplitz matrix  $T_N(f)$  in  $\text{Diag}(U_N)$  (see e.g. [9] or [27], pp. 80–88). In the following we will consider two directions. The first is algebraic and concerns the explicit construction of  $\mathcal{P}_{U_N}(T_N(f))$  in the four Hartley matrix algebras (see Subsection 7.1). The second direction is more analytic and concerns the approximation properties of  $\{\mathcal{P}_{U_N}(T_N(f))\}_N$  with respect to  $\{T_N(f)\}_N$  and its consequences in terms of the PCG method (see Subsection 7.2).

**7.1. Explicit formulas for optimal preconditioners.** The procedure for computing the Frobenius optimal approximation of a given symmetric Toeplitz matrix with real Fourier coefficients  $t_j = t_{|j|}$  ( $j \in \mathbf{Z}$ ) can be obtained by a simple minimization of the quadratic functional  $\mathcal{F}_{T_N(f)}(X) = \|T_N(f) - X\|_F^2$  over all the matrices  $X$  belonging to the algebra under consideration. In the case of the algebra of symmetric circulants, this leads to average formula for the T. Chan optimal preconditioner (see [10] or [27], pp. 56–61) described by  $\mathcal{P}_{F_N^1}(T_N(f)) = \text{circ}(\mathbf{a})$  where  $\mathbf{a} = (a_i)_{i=0}^{N-1}$  is  $J_N$ -even and

$$a_i = \frac{1}{N} ((N - i)t_i + it_{N-i}) \quad (i = 0, \dots, \lfloor N/2 \rfloor).$$

Therefore the arithmetic cost for determining the optimal preconditioner is proportional to the size  $N$ .

The same idea and the same kind of results hold for the case of Hartley matrix algebras for which the expression of the vectors  $\mathbf{a} = (a_i)_{i=0}^{N-1}$  and  $\mathbf{b} = (b_i)_{i=0}^{N-1}$  can be resumed as follows:

(7.2)

$$\begin{aligned} a_0 &= t_0, \\ a_i &= \frac{1}{N} ((N - i)t_i + i(-1)^{t^X[0]} t_{N-i}) \quad (i = 1, \dots, \lfloor (N - t^X[0])/2 \rfloor), \\ b_0 &= 0, \\ b_i &= \frac{1}{N} (t_i - (-1)^{t^X[0]} t_{N-i})(1 - t^X[1]) \quad (i = 1, \dots, \lfloor (N - 1 + t^X[0])/2 \rfloor). \end{aligned}$$

Here, with reference to Table 2 and Table 4.11 (where more information on the four Hartley matrix algebras are compactly reported), we have the  $J_N'$ -even vector  $\mathbf{a} =$

$H_N^I \mathbf{x}_+$  and the  $J'_N$ -odd vector  $\mathbf{b} = H_N^I \mathbf{x}_-$  in the case of the Hartley matrix algebras of type I and III, while we have the  $J''_N$ -odd vector  $\mathbf{a} = H_N^{II} \mathbf{x}_+$  and  $J''_N$ -even vector  $\mathbf{b} = H_N^{II} \mathbf{x}_-$  in the case of the Hartley matrix algebras of type II and IV. Then the optimal preconditioner of the symmetric Toeplitz matrix  $T_N(f)$  reads as follows

$$\mathcal{P}_{H_N^X}(T_N(f)) = \begin{cases} \text{circ}(\mathbf{a}) - J'_N \text{circ}(\mathbf{b}) & \text{for } X = \text{I}, \\ \text{scirc}(\mathbf{a}) + J''_N \text{scirc}(\mathbf{b}) & \text{for } X = \text{II}, \\ \text{circ}(\mathbf{a}) - J_N \text{circ}(\mathbf{b}) & \text{for } X = \text{III}, \\ \text{scirc}(\mathbf{a}) + J_N \text{scirc}(\mathbf{b}) & \text{for } X = \text{IV}. \end{cases}$$

Note that  $\mathcal{P}_{H_N^I}(T_N(f))$  can be found in [4] and [27], p. 87.

**7.2. Approximation results for optimal preconditioners.** We start by recalling some useful facts on  $\mathcal{P}_{U_N}$ .

LEMMA 7.1. [14, 34] *Let  $U_N \in \mathbf{C}^{N \times N}$  be an unitary matrix. Then for arbitrary  $A, B \in \mathbf{C}^{N \times N}$ , the operator  $\mathcal{P}_{U_N}$  has the following properties:*

1.  $\mathcal{P}_{U_N}(A) = U_N \sigma(U_N^* A U_N) U_N^*$ , where  $\sigma(X)$  is the diagonal matrix having  $(X)_{i,i}$  as diagonal elements.
2.  $\mathcal{P}_{U_N}(\alpha A + \beta B) = \alpha \mathcal{P}_{U_N}(A) + \beta \mathcal{P}_{U_N}(B)$  for all  $\alpha, \beta \in \mathbf{C}$ .
3.  $\mathcal{P}_{U_N}(A^*) = (\mathcal{P}_{U_N}(A))^*$ .
4.  $\text{tr}(\mathcal{P}_{U_N}(A)) = \text{tr}(A)$ .
5.  $\|\mathcal{P}_{U_N}\| = 1$  with the operator 2-norm  $\|\cdot\|$ .
6.  $\|\mathcal{P}_{U_N}\| = 1$  with the operator Frobenius norm  $\|\cdot\|_F$ .
7.  $\|A - \mathcal{P}_{U_N}(A)\|_F^2 = \|A\|_F^2 - \|\mathcal{P}_{U_N}(A)\|_F^2$  (law of Pythagoras).
8. If  $U_N = U_{N_1} \otimes \cdots \otimes U_{N_m}$  with unitary matrices  $U_{N_h}$  ( $h = 1, \dots, m$ ) and  $A_N = A_{N_1} \otimes \cdots \otimes A_{N_m}$ , then  $\mathcal{P}_{U_N}(A_N) = \mathcal{P}_{U_{N_1}}(A_{N_1}) \otimes \cdots \otimes \mathcal{P}_{U_{N_m}}(A_{N_m})$ .

In order to properly state the ‘‘matrix approximation results’’, we introduce the following concepts of ‘‘matrix convergence’’. We consider the sequence of matrix algebras  $\{\text{Diag}(U_N)\}_N$  of increasing  $N$  with the associated operators  $\mathcal{P}_{U_N}$ . We say that  $\{\mathcal{P}_{U_N}(T_N(f))\}_N$  converges strongly to  $\{T_N(f)\}_N$  for  $N \rightarrow \infty$ , if for any  $\epsilon > 0$ , there exists an index  $\bar{N}$  such that for all  $N \geq \bar{N}$ ,  $T_N(f) - \mathcal{P}_{U_N}(T_N(f))$  has eigenvalues in  $(-\epsilon, \epsilon)$  except for a constant number  $M_\epsilon$  of outliers (*proper clustering at zero*). Further we say that  $\{\mathcal{P}_{U_N}(T_N(f))\}_N$  converges weakly to  $\{T_N(f)\}_N$  for  $N \rightarrow \infty$ , if for any  $\epsilon > 0$ , there exists an index  $\bar{N}$  such that for all  $N \geq \bar{N}$ ,  $T_N(f) - \mathcal{P}_{U_N}(T_N(f))$  has eigenvalues in  $(-\epsilon, \epsilon)$  except for  $M_\epsilon = o(N)$  outliers (*general clustering at zero*). Furthermore, the convergence of  $\{\mathcal{P}_{U_N}(T_N(f))\}_N$  to  $\{T_N(f)\}_N$  for  $N \rightarrow \infty$  is called *k<sub>N</sub>-weak*, if  $M_\epsilon = O(k_N)$  and  $k_N = o(N)$  for  $N \rightarrow \infty$ . For brevity, we will omit to write ‘‘for  $N \rightarrow \infty$ ’’ when referring to matrix convergence.

In the case where we observe strong convergence, we say that the convergence is also *uniform*, if the number  $M_\epsilon$  does not depend on  $\epsilon$ . In the case where there is

strong convergence (strong or proper clustering in an alternative very popular terminology) and the function  $f$  is strictly positive, we have a superlinear convergence of the related PCG methods having  $\{\mathcal{P}_{U_N}(T_N(f))\}_N$  as preconditioner, but we may have a sublinear behaviour, if the weak (or general) convergence case occurs. Moreover, if the convergence is also uniform, that is  $M_\epsilon$  does not depend on  $\epsilon$ , the number of iterations initially decreases as the order  $N$  increases and then it stabilizes to a constant value. Therefore the associated PCG method is comparable with the one devised in [32].

The following Lemma 7.2 due to E. Tyrtshnikov provides a criterion to establish if convergence occurs, while Theorems 7.3, 7.4, and 7.5 reduce the analysis concerning the matrix approximation of Toeplitz matrices  $T_N(f)$  from the case of continuous  $2\pi$ -periodic symbols to the case of three trigonometric polynomials.

LEMMA 7.2. [38] *For all  $N \in \mathbf{N}$ , let hermitian matrices  $A_N, B_N \in \mathbf{C}^{N \times N}$  be given. If  $\|A_N - B_N\|_F^2 = O(1)$  for  $N \rightarrow \infty$ , then the sequence  $\{A_N\}_N$  converges strongly to  $\{B_N\}_N$ . If  $\|A_N - B_N\|_F^2 = o(N)$  for  $N \rightarrow \infty$ , then the convergence is weak.*

THEOREM 7.3. [34] *Let  $p$  be a trigonometric polynomial of fixed degree (independent of  $N$ ). If there exists an ordering of the eigenvalues  $\lambda_j$  ( $j = 0, \dots, N-1$ ) of  $\mathcal{P}_{U_N}(T_N(g))$  such that*

$$(7.3) \quad \limsup_{N \rightarrow \infty} N \cdot \max_{j=0, \dots, N-1} |\lambda_j - g(2\pi j/N)| < \infty$$

*for  $g \in \{1, \sin(x), \cos(x)\}$  (trigonometric Korovkin test), then  $\{\mathcal{P}_{U_N}(T_N(p))\}_N$  converges strongly to  $\{T_N(p)\}_N$ .*

*Moreover, if  $p$  is also even and if there exists an ordering of the eigenvalues  $\lambda_j$  ( $j = 0, \dots, N-1$ ) of  $\mathcal{P}_{U_N}(T_N(g))$  such that (7.3) is fulfilled for  $g \in \{1, \cos(x), \cos(2x)\}$  (even trigonometric Korovkin test), then  $\{\mathcal{P}_{U_N}(T_N(p))\}_N$  converges strongly to  $\{T_N(p)\}_N$ .*

THEOREM 7.4. *Let  $f : \mathbf{R} \rightarrow \mathbf{R}$  be a  $2\pi$ -periodic continuous function. If  $\{\mathcal{P}_{U_N}(T_N(p))\}_N$  converges strongly to  $\{T_N(p)\}_N$  for all the trigonometric polynomials  $p$  of fixed degree (independent of  $N$ ), then  $\{\mathcal{P}_{U_N}(T_N(f))\}_N$  converges strongly to  $\{T_N(f)\}_N$ .*

*Moreover, if  $f$  is also even and if  $\{\mathcal{P}_{U_N}(T_N(p))\}_N$  converges strongly to  $\{T_N(p)\}_N$  for all the even trigonometric polynomials  $p$  of fixed degree (independent of  $N$ ), then the sequence  $\{\mathcal{P}_{U_N}(T_N(f))\}_N$  converges strongly to  $\{T_N(f)\}_N$ .*

*Proof.* The proof of the first part can be found in Theorem 3.1 of [34]. The proof of the second part is a simple variation of the first one. The key is that the trigonometric polynomial of best  $L^\infty$  approximation of a  $2\pi$ -periodic continuous real-valued even function is a cosine polynomial, i.e. an even trigonometric polynomial (for this result see the elegant argument used in [23], p. 13).  $\square$

**THEOREM 7.5.** *Let  $f : \mathbf{R} \rightarrow \mathbf{R}$  be a  $2\pi$ -periodic continuous function and let  $k_N = o(N)$  for  $N \rightarrow \infty$ . If  $\{\mathcal{P}_{U_N}(T_N(p))\}_N$  converges  $k_N$ -weakly to  $\{T_N(p)\}_N$  for all the trigonometric polynomials  $p$  of fixed degree (independent of  $N$ ), then the sequence  $\{\mathcal{P}_{U_N}(T_N(f))\}_N$  converges  $k_N$ -weakly to  $\{T_N(f)\}_N$ .*

*Moreover, if  $f$  is also even and if  $\{\mathcal{P}_{U_N}(T_N(p))\}_N$  converges  $k_N$ -weakly to  $\{T_N(p)\}_N$  for all the even trigonometric polynomials  $p$  of fixed degree (independent of  $N$ ), then  $\{\mathcal{P}_{U_N}(T_N(f))\}_N$  converges  $k_N$ -weakly to  $\{T_N(f)\}_N$ .*

*Proof.* This is a variation of Theorem 3.2 in [34]. For the sake of completeness we report the whole proof. Let  $p_k$  be the trigonometric polynomial having degree  $k$  of best approximation of  $f$  in supremum norm [23]. For any  $\epsilon > 0$ , fix the integer  $M$  such that  $\|f - p_M\|_\infty < \epsilon/3$ . Then, by using a theorem of G. Szegö (see [18], p. 64) and Lemma 7.1, item 5, we have  $\|T_N(f) - T_N(p_M)\|_2 < \epsilon/3$  and  $\|\mathcal{P}_{U_N}(T_N(f)) - \mathcal{P}_{U_N}(T_N(p_M))\|_2 < \epsilon/3$ . Therefore, from the identity

$$\begin{aligned} T_N(f) - \mathcal{P}_{U_N}(T_N(f)) &= T_N(f) - T_N(p_M) - (\mathcal{P}_{U_N}(T_N(f)) - \mathcal{P}_{U_N}(T_N(p_M))) \\ &\quad + T_N(p_M) - \mathcal{P}_{U_N}(T_N(p_M)) \end{aligned}$$

we have that, except for a term of norm bounded by  $2\epsilon/3$ , the difference  $T_N(f) - \mathcal{P}_{U_N}(T_N(f))$  coincides with  $T_N(p_M) - \mathcal{P}_{U_N}(T_N(p_M))$ . From the hypothesis of  $k_N$ -weak convergence, we may split the hermitian matrix  $T_N(p_M) - \mathcal{P}_{U_N}(T_N(p_M))$  into two parts. The first part has a norm bounded by  $\epsilon/3$  and the second part has rank bounded by a universal constant times  $k_N$ . Therefore the claimed result is obtained, by invoking Cauchy's interlace theorem [40]. In the case, where  $f$  is also even, it is enough to observe that the trigonometric polynomial  $p_M$  in the latter argument can be chosen even too (see again [23], p. 13).  $\square$

Corollaries 7.6 and 7.7, trivial consequences of Theorems 7.4 and 7.5 respectively, are particularly useful for deriving and analyzing good preconditioners for the PCG method.

**COROLLARY 7.6.** *If the assumptions of Theorem 7.4 are fulfilled and if  $f$  is also positive, then for every  $\epsilon > 0$  and for all sufficiently large  $N$ , the matrix*

$$(\mathcal{P}_{U_N}(T_N(f)))^{-1}T_N(f)$$

*has eigenvalues in  $(1 - \epsilon, 1 + \epsilon)$  except  $M_\epsilon = O(1)$  outliers, at most.*

**COROLLARY 7.7.** *If the assumptions of Theorem 7.5 are fulfilled and if  $f$  is also positive, then for every  $\epsilon > 0$  and for all sufficiently large  $N$ , the matrix*

$$(\mathcal{P}_{U_N}(T_N(f)))^{-1}T_N(f)$$

*has eigenvalues in  $(1 - \epsilon, 1 + \epsilon)$  except  $M_\epsilon = O(k_N) = o(N)$  outliers, at most.*

We are now ready for analyzing the approximation properties of the Hartley matrix algebras with respect to Toeplitz structures.

**THEOREM 7.8.** *Let  $f : \mathbf{R} \rightarrow \mathbf{R}$  be an even  $2\pi$ -periodic continuous function. Then for every  $X \in \{\text{I, II, III, IV}\}$  the following facts hold:*

1.  $\{\mathcal{P}_{H_N^X}(T_N(f))\}_N$  converges strongly to  $\{T_N(f)\}_N$ .
2. If  $f$  is also positive, then for every  $\epsilon > 0$ , the matrices

$$(\mathcal{P}_{H_N^X}(T_N(f)))^{-1}T_N(f)$$

have eigenvalues in  $(1 - \epsilon, 1 + \epsilon)$  except  $M_\epsilon = O(1)$  outliers, at most.

*Proof.* By Theorem 7.4 we can reduce the proof of the first item to the convergence of the sequence  $\{\mathcal{P}_{H_N^X}(T_N(p))\}_N$  to  $\{T_N(p)\}_N$  for every even trigonometric polynomial  $p$ . Moreover, by invoking Theorem 7.3, the latter is true, if we prove that (7.3) is fulfilled for  $g \in \{1, \cos(x), \cos(2x)\}$ , where the  $\lambda_j$  ( $j = 0, \dots, N-1$ ), are the suitably ordered eigenvalues of  $\mathcal{P}_{U_N}(T_N(g))$ . For  $g(x) \equiv 1$  the desired result is trivial, since  $T_N(g) = I_N$  and therefore  $\mathcal{P}_{U_N}(T_N(g)) = I_N$ . The only nontrivial calculation is the explicit computation of the eigenvalues of  $\mathcal{P}_{U_N}(T_N(g))$  with  $g \in \{\cos(x), \cos(2x)\}$ . Indeed, looking at (7.3), we are interested in the eigenvalues  $\lambda_j(\mathcal{P}_{H_N^X}(T^{(s)}))$  of  $T^{(s)} := T_N(\cos(sx))$  ( $s = 1, 2; j = 0, \dots, N-1$ ). The entries  $\{t_k^{(s)}\}_{k \geq 0}$  of  $T^{(s)}$  are given by  $t_k^{(s)} = \frac{1}{2}\delta_{sk}$  and therefore the entries  $a_i^{(s)}, b_i^{(s)}$  of  $\mathcal{P}_{H_N^X}(T^{(s)})$ , for  $i$  as in (7.2), are identically zero except for  $a_s^{(s)} = \frac{N-s}{2N}$  and  $b_s^{(s)} = \frac{1}{2N}(1 - t^X[1])$ . For instance, for  $T^{(1)}$  we have  $\mathcal{P}_{H_N^X}(T^{(1)}) = H_N^X \text{diag}(\sqrt{N}\mathbf{x}) [H_N^X]^{-1}$  with  $\mathbf{x} = \mathbf{x}_+ + \mathbf{x}_-$  and

$$\mathbf{a}^{(1)} = \frac{N-1}{2N} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ (-1)^{t^X[0]} \end{pmatrix} = H_N^{I+t^X[0]} \mathbf{x}_+, \mathbf{b}^{(1)} = \frac{1-t^X[1]}{2N} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ (-1)^{1+t^X[0]} \end{pmatrix} = H_N^{I+t^X[0]} \mathbf{x}_-,$$

while for  $T^{(2)}$  we find

$$\mathbf{a}^{(2)} = \frac{N-1}{2N} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ (-1)^{t^X[0]} \end{pmatrix} = H_N^{I+t^X[0]} \mathbf{x}_+, \mathbf{b}^{(2)} = \frac{1-t^X[1]}{2N} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ (-1)^{1+t^X[0]} \end{pmatrix} = H_N^{I+t^X[0]} \mathbf{x}_-.$$

In general, for  $j = 0, \dots, N-1$ , we find

$$\begin{aligned} \lambda_j(\mathcal{P}_{H_N^X}(T^{(s)})) &= (\sqrt{N}\mathbf{x})_j = (\sqrt{N}H_N^{I+2t^X[0]}(\mathbf{a}^{(s)} + \mathbf{b}^{(s)}))_j \\ &= (\sqrt{N}H_N^{I+2t^X[0]}\mathbf{a}^{(s)})_j + (1 - t^X[1])O\left(\frac{1}{N}\right) \end{aligned}$$

since  $\mathbf{b}^{(s)}$  is infinitesimal and indeed, for  $j = 0, \dots, N-1$ , we have

$$|(\sqrt{N}H_N^{I+2t^X[0]}\mathbf{b}^{(s)})_j| = |\text{cas}(\cdot)b_s^{(s)} \pm \text{cas}(\cdot)b_{N-s}^{(s)}| \leq (1 - t^X[1])\frac{\sqrt{2}}{N}.$$

Since  $t^{1+2t^X[0]}[0] = 0$  and  $t^{1+2t^X[0]}[1] = t^X[0]$ , the Hartley matrix  $H_N^{1+2t^X[0]}$  (from the representation in  $\mathbf{a}^{(s)}$  to the representation in  $\lambda$ ) is

$$H_N^{1+2t^X[0]} = \frac{1}{\sqrt{N}} \left( \text{cas} \left( \frac{\pi(2i+t^X[0])j}{N} \right) \right)_{i,j=0}^{N-1}.$$

For  $j = 0, \dots, N-1$ , we deduce

$$\begin{aligned} (\sqrt{N}H_N^{1+2t^X[0]}\mathbf{a}^{(s)})_j &= \frac{N-s}{2N} \left[ \text{cas} \left( \frac{\pi(2j+t^X[0])s}{N} \right) + (-1)^{t^X[0]} \text{cas} \left( \frac{\pi(2j+t^X[0])(N-s)}{N} \right) \right] \\ &= \frac{N-s}{N} \cos \left( s \frac{\pi(2j+t^X[0])}{N} \right) \end{aligned}$$

and as a consequence

$$\lambda_j(\mathcal{P}_{H_N^X}(T^{(s)})) = \frac{N-s}{N} \cos \left( s \frac{\pi(2j+t^X[0])}{N} \right) + (1-t^X[1])O\left(\frac{1}{N}\right)$$

where, more precisely, the term  $O(N^{-1})$  has the form

$$\frac{1}{N} \sin \left( s \frac{\pi(2j+t^X[0])}{N} \right).$$

In conclusion, for  $j = 0, \dots, N-1$

$$\lambda_j(\mathcal{P}_{H_N^X}(T^{(s)})) = \underbrace{\frac{N-s}{N} \cos \left( s \frac{\pi(2j+t^X[0])}{N} \right)}_{\text{what we wanted}} + \frac{(1-t^X[1])}{N} \sin \left( s \frac{\pi(2j+t^X[0])}{N} \right).$$

Hence, setting

$$\phi_i^X(x) = \frac{1}{\sqrt{N}} \text{cas} \left( \frac{2i+t^X[1]}{2} x \right), \quad x_j^X = \frac{(2j+t^X[0])\pi}{N},$$

we find that  $H_N^X$  is a Gram matrix of the form  $H_N^X = (\phi_i^X(x_j^X))_{i,j=0}^{N-1}$ . Therefore we have compactly proved

$$\lambda_j(\mathcal{P}_{H_N^X}(T^{(s)})) = \frac{N-s}{N} \cos(s x_j^X) + (1-t^X[1])O\left(\frac{1}{N}\right)$$

which clearly implies (7.3) with  $g \in \{1, \cos(x), \cos(2x)\}$ . Furthermore the proof of the second item is a direct consequence of the first part and of Corollary 7.6.  $\square$

Finally mention has to be made to the case of complex-valued even symbols. In this setting it is easy to prove that  $T_N(f)$  has complex entries: however it is easy to see that the sequence  $\{T_N(|f|^2)\}_N$  converges strongly to  $\{T_N^*(f)T_N(f)\}_N$  for  $N \rightarrow \infty$  and, by the previous results,  $\{\mathcal{P}_{U_N}(T_N(|f|^2))\}_N$  converges strongly to  $\{T_N(|f|^2)\}_N$  for  $N \rightarrow \infty$  since  $|f|^2$  is continuous real-valued and even. Therefore with an additional little effort it can be proven that both  $\{\mathcal{P}_{U_N}(T_N(|f|^2))\}_N$  and  $\{\mathcal{P}_{U_N}(T_N^*(f)T_N(f))\}_N$  converges strongly to  $\{T_N^*(f)T_N(f)\}_N$  for  $N \rightarrow \infty$  for  $f$  simply continuous. We notice that the latter claim extends a result by Potts and Steidl where the assumption was stronger and concerned the Wiener class (see [29]). However we should also notice that, in this context, unless one has the normal equation system to solve, a better idea is to use first a preconditioner (for instance in the circulant algebra) and then to pass to the normal equations for the preconditioned system.

**8. Multilevel Toeplitz matrices.** By following the notations of E. Tyrtyshnikov, a *multilevel Toeplitz matrix of level  $m$  and dimension  $N_1 \times \cdots \times N_m$*  is defined as the matrix generated by the Fourier coefficients of a  $m$ -variate Lebesgue integrable function  $f : (-\pi, \pi]^m \rightarrow \mathbf{R}$  according to the law given in equations (6.1) at p. 23 of [38]. In our context, according to the one-level case, we are interested in multilevel Toeplitz matrices which are symmetric at every level. Therefore we have  $f(x_1, \dots, x_m) = f(|x_1|, \dots, |x_m|)$  on  $(-\pi, \pi]^m$  and then the Fourier coefficients forming the symmetric  $m$ -level Toeplitz matrix  $T_N(f)$  fulfil

$$t_j = t_{|j|} \quad (j := (j_1, \dots, j_m) \in \{1 - N_1, \dots, N_1 - 1\} \times \cdots \times \{1 - N_m, \dots, N_m - 1\})$$

with  $|j| := (|j_1|, \dots, |j_m|)$ . Similarly, for given unitary matrices  $U_{N_h}$  ( $h = 1, \dots, m$ ) related to transforms of one-level algebras, the corresponding  *$m$ -level matrix algebra* is defined as the set of all matrices simultaneously diagonalized by means of the following Kronecker product

$$(8.1) \quad U_N = U_{N_1} \otimes \cdots \otimes U_{N_m} \quad (N := (N_1, \dots, N_m)).$$

For instance the  $m$ -level Hartley matrix algebra of type I is defined as in (8.1) with  $U_{N_h} = H_{N_h}^I$  ( $h = 1, \dots, m$ ). As in the one-level case, the analysis is done in two directions, first the study of the algebraic and computational features of optimal preconditioners and then approximation properties.

**8.1. Explicit formulas for optimal preconditioners in the multilevel case.** Concerning the multilevel case we have a canonical way for giving an explicit formula for the optimal preconditioner. The idea is purely algebraic and is based on the multi-index notation (while does not refer explicitly to the symbol). Indeed if  $m \geq 2$  and  $N = (N_1, \dots, N_m)$  is a multi-index, then a corresponding real  $m$ -level Toeplitz matrix  $T_N$  (with  $m$  levels of symmetry) can be described recursively as

$$T_N = \begin{pmatrix} \tilde{t}_0 & \tilde{t}_1 & \cdots & \tilde{t}_{N_1-2} & \tilde{t}_{N_1-1} \\ \tilde{t}_1 & \ddots & \ddots & & \tilde{t}_{N_1-2} \\ \vdots & \ddots & & \ddots & \vdots \\ \tilde{t}_{N_1-2} & & \ddots & \ddots & \tilde{t}_1 \\ \tilde{t}_{N_1-1} & \tilde{t}_{N_1-2} & \cdots & \tilde{t}_1 & \tilde{t}_0 \end{pmatrix}$$

where every  $\tilde{t}_j$  is a real  $(m-1)$ -level Toeplitz matrix (with  $m-1$  levels of symmetry) and a real one-level matrix with one level of symmetry is a standard real symmetric matrix. Correspondingly, let  $t_j := t_j^X$  be the optimal preconditioner of level  $m-1$  of the single block  $\tilde{t}_j$  (with respect to the  $(m-1)$ -level Hartley matrix algebra of type  $X$ ), then the optimal preconditioner of  $T_N$  (with respect to the  $m$ -level Hartley matrix

algebra of type  $X$ ) is obtained by applying formally the rule (7.2) to the matrix

$$Z_N = \begin{pmatrix} t_0 & t_1 & \cdots & t_{N_1-2} & t_{N_1-1} \\ t_1 & \ddots & \ddots & & t_{N_1-2} \\ \vdots & \ddots & & \ddots & \vdots \\ t_{N_1-2} & & \ddots & \ddots & t_1 \\ t_{N_1-1} & t_{N_1-2} & \cdots & t_1 & t_0 \end{pmatrix}.$$

**8.2. Approximation results for optimal preconditioners in the multi-level case.** We are interested in extending the results proven in the Section 7.2 to  $m$  dimensions. In fact, we analyze what is necessary to have and, especially, what is kept when we switch from one dimension to  $m$  dimensions. Surprisingly enough, we find that all the used tools hold or have a version in  $m$  dimensions: for instance Lemma 7.2, Theorems 7.4, 7.5, and Corollaries 7.6, 7.7 contain statements not depending on the structure of the matrices and item 7 of Lemma 7.1 is valid for any matrix algebra and so for multilevel matrix algebras as well (recall that  $U_N$  in (8.1) is unitary). The only inherently one-level result is Theorem 7.3 for which we state the following  $m$ -dimensional version.

**THEOREM 8.1.** [34] *Let  $p$  be an  $m$ -variate trigonometric polynomial of fixed degree (independent of the multi-index  $N = (N_1, \dots, N_m)$ ). Further let*

$$(8.2) \quad k_N := \left( \prod_{j=1}^m N_j \right) \left( \sum_{j=1}^m N_j^{-1} \right).$$

and  $j/N := (j_1/N_1, \dots, j_m/N_m)$  for  $j = (j_1, \dots, j_m)$ . If there exists an ordering of the eigenvalues  $\lambda_j$  (with  $j = (j_1, \dots, j_m)$  and  $j_k = 0, \dots, N_k - 1$  for  $k = 1, \dots, m$ ) of  $\mathcal{P}_{U_N}(T_N(g))$  such that for every  $k = 1, \dots, m$  it holds

$$(8.3) \quad \limsup_{N_k \rightarrow \infty} N_k \cdot \max_{\substack{j = (j_1, \dots, j_m), \\ j_q = 0, \dots, N_q - 1, \\ 1 \leq q \leq m}} |\lambda_j - g(2\pi j/N)| < \infty,$$

for  $g \in \{1, \sin(x_1), \cos(x_1), \dots, \sin(x_m), \cos(x_m)\}$ , then  $\{\mathcal{P}_{U_N}(T_N(p))\}_N$  converges  $k_N$ -weakly to  $\{T_N(p)\}_N$  ( $m$ -level trigonometric Korovkin test).

Moreover, if  $p$  is also even with respect to every variable, i.e.  $p(x_1, \dots, x_m) = p(|x_1|, \dots, |x_m|)$ , and if there exists an ordering of the eigenvalues  $\lambda_j$  ( $j = (j_1, \dots, j_m)$  ( $j_k = 1, \dots, N_k; 1 \leq k \leq m$ )) of  $\mathcal{P}_{U_N}(T_N(g))$  such that for every  $k = 1, \dots, m$  the condition (8.3) is fulfilled for

$$g \in \{1, \cos(x_1), \cos(2x_1), \dots, \cos(x_m), \cos(2x_m)\},$$

then  $\{\mathcal{P}_{U_N}(T_N(p))\}_N$  converges  $k_N$ -weakly to  $\{T_N(p)\}_N$  ( $m$ -level even trigonometric Korovkin test).

Therefore, we instantly deduce the validity in  $m$  dimensions of a generalization of Theorem 7.8.

**THEOREM 8.2.** *Let  $f : \mathbf{R}^m \rightarrow \mathbf{R}$  be an  $m$ -variate even  $2\pi$ -periodic continuous function. Let  $N = (N_1, \dots, N_m)$  with  $m \geq 2$  be a multi-index and let  $k_N$  be defined by (8.2). Then for every  $X \in \{\text{I, II, III, IV}\}$  the following facts hold:*

1.  $\{\mathcal{P}_{H_N^X}(T_N(f))\}_N$  and  $\{\mathcal{P}_{F_N^X}(T_N(f))\}_N$  converges  $k_N$ -weakly to  $\{T_N(f)\}_N$ .
2. If  $f$  is also positive, then for every  $\epsilon > 0$ , the matrices

$$(\mathcal{P}_{H_N^X}(T_N(f)))^{-1}T_N(f)$$

have eigenvalues in  $(1 - \epsilon, 1 + \epsilon)$  except  $M_\epsilon = O(k_N)$  outliers, at most.

*Proof.* By Theorem 7.5 we can reduce the proof of the first item to the  $k_N$ -weak convergence of the sequence  $\{\mathcal{P}_{H_N^X}(T_N(p))\}_N$  to  $\{T_N(p)\}_N$  for every even trigonometric polynomial  $p$ . Furthermore, by Theorem 8.1, the latter is true if we prove that the conditions (8.3) are fulfilled for every  $k = 1, \dots, m$  and for  $g \in \{1, \cos(x_1), \cos(2x_1), \dots, \cos(x_m), \cos(2x_m)\}$ , where  $\lambda_j$  ( $j = 0, \dots, N_1 N_2 \dots N_m - 1$ ) are suitably ordered eigenvalues of  $\mathcal{P}_{U_N}(T_N(g))$ . For  $g(x) \equiv 1$ , the desired result is trivial, since  $T_N(g) = I_N$  and therefore  $\mathcal{P}_{U_N}(T_N(g)) = I_N$ . Now it is enough to observe that  $T_N(\cos(jx_k)) = I_{N_1} \otimes \dots \otimes I_{N_{k-1}} \otimes T_{N_k}(\cos(jx_k)) \otimes I_{N_{k+1}} \otimes \dots \otimes I_{N_m}$  ( $j = 1, 2; k = 1, \dots, m$ ), and therefore by (8.1) and Lemma 7.1, item 8, we have

$$\mathcal{P}_{U_N}(T_N(\cos(jx_k))) = I_{N_1} \otimes \dots \otimes I_{N_{k-1}} \otimes \mathcal{P}_{U_{N_k}}(T_{N_k}(\cos(jx_k))) \otimes I_{N_{k+1}} \otimes \dots \otimes I_{N_m}.$$

The above identities imply that (7.3) in the one-dimensional case are sufficient to conclude that the conditions (8.3) are fulfilled (and the first item is proven). Finally the proof of the second item is a direct consequence of the first part and of Corollary 7.7.  $\square$

We observe that the above results agree with a known fact in the two-level circulant and  $\tau$  cases (for the  $\tau$  matrix algebras see [3]). Indeed only the weak convergence has been proved because the number of the outliers is, in both cases, equal to  $O(n_1 + n_2)$  [8, 12] even if the function  $f$  is a bivariate trigonometric polynomial. More precisely, this means that the hypotheses of Theorem 7.4, regarding the strong approximation in the polynomial case, are not fulfilled by the two-level circulant and  $\tau$  algebras and therefore strong convergence cannot be proved in the general case. As a matter of fact, in [35] and [36] it has been proved that any sequence of preconditioners belonging to ‘‘partially equimodular’’ algebras [36] cannot be superlinear (i.e. the approximation cannot be strong) for sequence of multilevel Toeplitz matrices generated by simple positive polynomials. Here, ‘‘partially equimodular’’ refers to some very weak assumptions on  $U_N$  that are instantly fulfilled by all the known multilevel trigonometric matrix algebras (circulants and Hartley matrix algebras included). Therefore the results reported in Theorem 8.2 which are not very satisfactory

for large  $m$  are however asymptotically the best that we can obtain when dealing with multilevel Toeplitz structures.

**8.3. Numerical experiments.** Here we just give a numerical evidence of the strong clustering properties of Toeplitz matrices preconditioned by Hartley matrices: more precisely we check it through the PCG algorithm when applied to symmetric Toeplitz systems with the optimal preconditioners described in Section 7.

As an example, we choose as generating function  $f(x) = x^2 + 1$ . Its Fourier coefficients are  $t_0 = 1 + \pi^2/3$  and  $t_k = (-1)^k \cdot 2/k^2$  for  $k \neq 0$ . We tried the PCG method for the coefficient matrix  $T_N(f)$ ,  $N = 2^p$ ,  $p = 4, \dots, 9$ . The known data vector is  $\mathbf{b}$  chosen in such a way that the solution coincides with the vector  $(k^{-1} + 1)_{k=3}^{N+2}$ .

Below is our notation:

- $T_{N,\sigma}$  indicates the matrix  $T_N = T_N(f)$  cut in such a way that  $t_k = 0$  if  $|k| > \sigma$ ;
- $\text{Chan}_N$  indicates the optimal circulant preconditioner (the one introduced by T. Chan) i.e.  $\mathcal{P}_{F_N^I}(T_N(f))$ ;
- $\tau_N$  denotes the optimal preconditioner in the  $\tau$  algebra.

$\log_2(N)$	4	5	6	7	8	9
$\text{Chan}_N$	8	8	7	7	7	6
$\tau_N$	6	6	5	5	5	4
$\mathcal{P}_{H_N^I}(T_N(f))$	9	8	7	7	7	6
$\mathcal{P}_{H_N^{II}}(T_N(f))$	9	8	7	7	7	6
$\mathcal{P}_{H_N^{III}}(T_N(f))$	8	8	7	7	7	6
$\mathcal{P}_{H_N^{IV}}(T_N(f))$	8	8	7	7	6	6
$T_{N,\log_2(N)}$	6	6	6	5	5	5
$T_{N, \lceil \sqrt{\log_2(N)} \rceil}$	8	8	8	8	8	8

We stop the iterations, if  $\|\mathbf{r}\|_2 \leq 10^{-9}\|\mathbf{b}\|_2$  with the usual residual  $\mathbf{r} := \mathbf{b} - T_N(f)\mathbf{x}$ . It is interesting to observe that all the preconditioners are substantially equivalent including the ones of band type with slowly increasing bandwidth: the  $\tau$  algebra optimal preconditioner and the band type preconditioner  $T_{N,\log_2(N)}$  behave slightly better than the others and this agrees with the analysis given in [33]. We remark that the study of these almost negligible differences between the algebras in the positive case is the subject of the paper [16]. Indeed in [33] it was proven that we have not to expect big differences among the performances of the different preconditioners when the continuous generating function is strictly positive, while, according to the analysis in [13], the difference could be remarkable in the nonnegative case (i.e. when the symbol has zeros): however, in the nonnegative case, again in [13], it is proven

that the main ingredient for a fast convergence is not the chosen algebra but the functional approximation process which defines the eigenvalues of the preconditioner. Since the Frobenius optimal approximations are related to linear positive operators (see [26]) and essentially to the Cesaro sum we observe a very slow convergence speed: this observation tells one that other kind of preconditioners (of Strang-type mainly, see the work by T. Kailath and V. Olshevsky [24, 25]) have to be preferred in the nonnegative case. Finally, we observe that the clustering analysis in [25] can be improved in the following two directions: the assumption on the Wiener class is not necessary for T. Chan-type preconditioners since only continuity is essential (see Theorem 7.8 and [34, 14] where the analysis is reduced to the Weierstrass theorem through the Korovkin theorem); the assumption on the Wiener class can be replaced by the hypothesis that the symbol belongs to the union of the Wiener class and the Dini-Lipschitz class (see [33] where the analysis is reduced to the convergence of the underlying approximation process through the notion of “good algebras”). For the definition of function spaces such as the Wiener algebra and the Dini-Lipschitz class see e.g. [41].

**9. Conclusions and remarks.** We have analyzed the DHTs of types I – IV and the related Hartley matrix algebras by proving representation formulas and orthogonal decompositions which involve circulants, skew-circulants and special flip-like matrices. We have proven that any of these DHTs of length  $N = 2^t$  can be factorized by means of a divide-and-conquer strategy into a product of sparse, orthogonal matrices where in this context sparse means at most two nonzero entries per row and column. The sparsity joint with orthogonality of the matrix factors is a key property which has been exploited for proving that these new algorithms have low arithmetic costs equal to  $\frac{5}{2}N \log_2(N) + O(N)$  arithmetic operations and an excellent normwise numerical stability. Furthermore, we have considered the best Frobenius approximation of a given symmetric Toeplitz matrix generated by an Lebesgue integrable symbol in the Hartley matrix algebra. We provided explicit formulas for computing optimal preconditioners of symmetric Toeplitz matrices, where the arithmetic cost are proportional to the size of the involved matrices. By using the matrix approximation theory, in the one-level setting we proved the strong clustering at unity of the preconditioned matrix sequences under the sole assumption of continuity and positivity of the generating function. The multilevel case is inherently more difficult and this agrees with the negative results obtained by the second author and E. Tyrtyshnikov. Finally a future work (the preliminary results are very encouraging) should concern a careful implementation of the fast DHT algorithms proposed in this paper in order check how tight are the worst case bounds derived in Section 6 for the numerical stability (see [1]).

## REFERENCES

- [1] A. ARICO', S. SERRA CAPIZZANO, AND M. TASCHE, *Numerical stability of some fast DHT algorithms*, in preparation.
- [2] G. BI AND Y. CHEN, *Fast generalized DFT and DHT algorithms*, *Signal Process.*, 65(1998), pp. 383–390.
- [3] D. BINI AND M. CAPOVANI, *Spectral and computational properties of band symmetric Toeplitz matrices*, *Linear Algebra Appl.*, 52/53(1983), pp. 99–126.
- [4] D. BINI AND P. FAVATI, *On a matrix algebra related to the discrete Hartley transform*, *SIAM J. Matrix Anal. Appl.*, 14-2(1993), pp. 500–507.
- [5] A. BORTOLETTI AND C. DI FIORE, *On a set of matrix algebras related to discrete Hartley-type transforms*, *Linear Algebra Appl.*, 366(2003), pp. 65–85.
- [6] R. N. BRACEWELL, *The Hartley Transform*, Oxford Univ. Press, Oxford, 1986.
- [7] V. BRITANAK AND K.R. RAO, *The fast generalized discrete Fourier transforms: A unified approach to the discrete sinusoidal transforms computation*, *Signal Process.*, 79-2(1999), pp. 135–150.
- [8] R. H. CHAN AND X. JIN, *A family of block preconditioners for block systems*, *SIAM J. Sci. Stat. Comput.*, 13(1992), pp. 1218–1235.
- [9] R. H. CHAN, M. K. NG, *Conjugate gradient methods of Toeplitz systems*, *SIAM Rev.*, 38(1996), pp. 427–482.
- [10] T. F. CHAN, *An optimal circulant preconditioner for Toeplitz systems*, *SIAM J. Sci. Stat. Comp.*, 9(1988), pp. 766–771.
- [11] P. J. DAVIS, *Circulant Matrices*, Wiley, New York, 1979.
- [12] F. DI BENEDETTO, *Preconditioning of block Toeplitz matrices by sine transforms*, *SIAM J. Sci. Comput.*, 18(1997), pp. 499–515.
- [13] F. DI BENEDETTO AND S. SERRA CAPIZZANO, *A unifying approach to abstract matrix algebra preconditioning*, *Numer. Math.*, 82-1(1999), pp. 57–90.
- [14] F. DI BENEDETTO AND S. SERRA CAPIZZANO, *Optimal multilevel matrix algebra operators*, *Linear Multilinear Algebra*, 48(2000), pp. 35–66.
- [15] C. DI FIORE, *Matrix algebras and displacement decompositions*, *SIAM J. Matrix Anal. Appl.*, 21-2(2000), pp. 646–667.
- [16] C. DI FIORE, S. FANELLI, AND P. ZELLINI, *On the best least squares fit to a matrix and its applications*, *Advances in Mathematics*, Nova Science Publ.
- [17] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1983.
- [18] U. GRENANDER AND G. SZEGÖ, *Toeplitz Forms and Their Applications*, Second Edition, Chelsea, New York, 1984.
- [19] G. HEINIG AND K. ROST, *Hartley transform representations of symmetric Toeplitz matrix inverses with application to fast matrix-vector multiplication*, *SIAM J. Matrix Anal. Appl.*, 22-1(2000), pp. 86–105.
- [20] G. HEINIG AND K. ROST, *Hartley transform representations of inverses of real Toeplitz-plus-Hankel matrices*, *Numer. Funct. Anal. Optimization*, 21-1/2(2000), pp. 175–189.
- [21] N. J. HIGHAM, *Accuracy and Stability of Numerical Algorithms*, 2nd ed, SIAM, Philadelphia, 2002.
- [22] N. C. HU, H. I. CHANG, AND O. K. ERSOY, *Generalized discrete Hartley transforms*, *IEEE Trans. Signal Process.*, 40-12(1992), pp. 2931–2940.
- [23] D. JACKSON, *The Theory of Approximation*, Amer. Math. Soc., New York, 1930.
- [24] T. KAILATH AND V. OLSHEVSKY, *Displacement structure approach to discrete-trigonometric-transform based preconditioners of G. Strang type and T. Chan type*, *Calcolo*, 33(1996), pp. 191–208.

- [25] T. KAILATH AND V. OLSHEVSKY, *Displacement structure approach to discrete-trigonometric-transform based preconditioners of G. Strang type and T. Chan type*, SIAM J. Matrix Anal. Appl., 26-3(2005), pp. 706–734.
- [26] P. P. KOROVKIN, *Linear Operators and Approximation Theory* (English translation), Hindustan Publishing Co., Delhi, 1960.
- [27] M. K. NG, *Iterative Methods for Toeplitz Systems*, Oxford Univ. Press, Oxford, 2004.
- [28] G. PLONKA AND M. TASCHE, *Fast and numerically stable algorithms for discrete cosine transforms*, Linear Algebra Appl., 394(2005), pp. 309–345.
- [29] D. POTTS AND G. STEIDL, *Optimal trigonometric preconditioners for nonsymmetric Toeplitz systems*, Linear Algebra Appl., 281(1998), pp. 265–292.
- [30] K. R. RAO AND P. YIP, *Discrete Cosine Transform: Algorithms, Advantages, and Applications*, Academic Press, Boston, 1990.
- [31] U. SCHREIBER, *Fast and numerically stable trigonometric transforms* (in German), Thesis, Univ. of Rostock, 1999.
- [32] S. SERRA, *Optimal, quasi-optimal and superlinear band-Toeplitz preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems*, Math. Comput., 66(1997), pp. 651–665.
- [33] S. SERRA CAPIZZANO, *Toeplitz preconditioners constructed from linear approximation processes*, SIAM J. Matrix Anal. Appl., 20-2(1998), pp. 446–465.
- [34] S. SERRA CAPIZZANO, *A Korovkin-type theory for finite Toeplitz operators via matrix algebras*, Numer. Math., 82-1(1999), pp. 117–142.
- [35] S. SERRA CAPIZZANO AND E. TYRTYSHNIKOV, *Any circulant-like preconditioner for multi-level matrices is not superlinear*, SIAM J. Matrix Anal. Appl., 22-1(1999), pp. 431–439.
- [36] S. SERRA CAPIZZANO AND E. TYRTYSHNIKOV, *How to prove that a preconditioner can not be superlinear*, Math. Comput., 72(2003), pp. 1305–1316.
- [37] M. TASCHE AND H. ZEUNER, *Roundoff error analysis for fast trigonometric transforms*, in: Handbook of Analytic-Computational Methods in Applied Mathematics, G.A. Anastassiou, ed., CRC Press, Boca Raton, 2000, pp. 357–406.
- [38] E. TYRTYSHNIKOV, *A unifying approach to some old and new theorems on distribution and clustering*, Linear Algebra Appl., 232(1996), pp. 1–43.
- [39] C. F. VAN LOAN, *Computational Frameworks for the Fast Fourier Transform*, SIAM, Philadelphia, 1992.
- [40] J. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.
- [41] A. ZYGMUND, *Trigonometric Series*, Cambridge University Press, Cambridge, 1959.