2013

TR-2013010: Transformations of Matrix Structures Work Again II

Victor Y. Pan

Transformations of Matrix Structures Work Again II *

Victor Y. Pan[1,2],[a]

[1] Department of Mathematics and Computer Science
Lehman College of the City University of New York
Bronx, NY 10468 USA
[2] Ph.D. Programs in Mathematics and Computer Science
The Graduate Center of the City University of New York
New York, NY 10036 USA
[a] victor.pan@lehman.cuny.edu
http://comet.lehman.cuny.edu/vpan/

Abstract

Matrices with the structures of Toeplitz, Hankel, Vandermonde and Cauchy types are omnipresent in modern computations in Sciences, Engineering and Signal and Image Processing. The four matrix classes have distinct features, but in [P90] we showed that Vandermonde and Hankel multipliers transform all these structures into each other and proposed to employ this property in order to extend any successful algorithm that inverts matrices of one of these four classes to inverting matrices with the structures of the three other types. The power of this approach was widely recognized later, when novel numerically stable algorithms solved nonsingular Toeplitz linear systems of equations in quadratic (versus classical cubic) arithmetic time based on transforming Toeplitz into Cauchy matrix structures. More recent papers combined such a transformation with a link of the Cauchy matrices to the Hierarchical Semiseparable matrix structure, which is a specialization of matrix representations employed by the Fast Multipole Method. This produced numerically stable algorithms that approximated the solution of a nonsingular Toeplitz linear system of equations in nearly linear arithmetic time. We first revisit the successful method of structure transformation, covering it comprehensively. Then we analyze the latter efficient approximation algorithms for Toeplitz linear systems and extend them to approximate the products of Vandermonde and Cauchy matrices by a vector and the solutions of Vandermonde and Cauchy linear systems of equations where they are nonsingular and well conditioned. We decrease the arithmetic cost of the known numerical approximation algorithms for these tasks from quadratic to nearly linear, and similarly for the computations with the matrices of a more general class having structures of Vandermonde and Cauchy types and for polynomial and rational evaluation and interpolation. We also accelerate a little further the known numerical approximation algorithms for a nonsingular Toeplitz or Toeplitz-like linear system by employing distinct transformations of matrix structures, and we briefly discuss some natural research challenges, particularly some promising applications of our techniques to high precision computations.

Keywords: Transformations of matrix structures, Vandermonde matrices, Cauchy matrices, Multipole method, HSS matrices, Toeplitz matrices

AMS Subject Classification: 15A04, 15A06, 15A09, 47A65, 65D05, 65F05, 68Q25

*Some results of this paper have been presented at the 18th Conference of the International Linear Algebra Society (ILAS'2013), Providence, RI, 2013 and at the 15th Annual Conference on Computer Algebra in Scientific Computing (CASC 2003), September 9–13, 2013, Berlin, Germany

1 Introduction

Table 1: Four classes of structured matrices

.

Toeplitz matrices $T=(t_{i-j})_{i,j=0}^{n-1}$	Hankel matrices $H=(h_{i+j})_{i,j=0}^{n-1}$
$\begin{pmatrix} t_0 & t_{-1} & \cdots & t_{1-n} \\ t_1 & t_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{-1} \\ t_{n-1} & \cdots & t_1 & t_0 \end{pmatrix}$	$\begin{pmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \iddots & h_n \\ \vdots & \iddots & \iddots & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n-2} \end{pmatrix}$
Vandermonde matrices $V=V_{\mathbf{s}}=\left(s_i^j\right)_{i,j=0}^{n-1}$	Cauchy matrices $C=C_{\mathbf{s},\mathbf{t}}=\left(\frac{1}{s_i-t_j}\right)_{i,j=0}^{n-1}$
$\begin{pmatrix} 1 & s_1 & \cdots & s_1^{n-1} \\ 1 & s_2 & \cdots & s_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & s_n & \cdots & s_n^{n-1} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{s_1-t_1} & \cdots & \frac{1}{s_1-t_n} \\ \frac{1}{s_2-t_1} & \cdots & \frac{1}{s_2-t_n} \\ \vdots & & \vdots \\ \frac{1}{s_n-t_1} & \cdots & \frac{1}{s_n-t_n} \end{pmatrix}$

Table 1 displays four classes of most popular structured matrices, which are omnipresent in modern computations for Sciences, Engineering, and Signal and Image Processing and which have been naturally extended to larger classes of matrices, $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, and $\mathcal{C}$, having structures of Toeplitz, Hankel, Vandermonde and Cauchy types, respectively. Such matrices can be readily expressed via their displacements of small ranks, which implies their further attractive properties:

- Compressed representation of a matrices as well as their products and inverses through a small number of parameters
- Multiplication by vectors in nearly linear arithmetic time
- Solution of nonsingular linear systems of equations with these matrices in quadratic or nearly linear arithmetic time

Highly successful research and implementation work based on these properties has been continuing for more than three decades. We follow [P90] and employ structured matrix multiplications to transform the four matrix structures into each other. For example, $\mathcal{T}\mathcal{H}=\mathcal{H}\mathcal{T}=\mathcal{H}$, $\mathcal{H}\mathcal{H}=\mathcal{T}$, whereas V^TV is a Hankel matrix. The paper [P90] showed that *such techniques enable one to extend any successful algorithm for the inversion of the matrices of any of the four classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, and $\mathcal{C}$ to the matrices of the three other classes, and similarly for the solution of linear systems of equations.* We cover this approach comprehensively and simplify its presentation versus [P90] because we employ Sylvester's displacements $AM-MB$, versus Stein's displacements $M-AMB$ in [P90], and apply the machinery of operating with them from [P00] and [P01, Section 1.5]. We study quite simple structure transforms, but they have surprising power where they link together matrix classes having distinct features.

For example, row and column interchanges destroy Toeplitz and Hankel but not Cauchy matrix structure, and [GKO95] and [G98] obtained numerically stable solution of Toeplitz and Hankel linear systems without pivoting by transforming the inputs into Cauchy-like matrices. The resulting algorithms run in quadratic arithmetic time versus classical cubic. Like [GKO95] and [G98] the papers [MRT05], [CGS07], [XXG12], and [XXCB] reduced the solution of a nonsingular Toeplitz linear system of n equations to computations with a special Cauchy matrix $C=C_{\mathbf{s},\mathbf{t}}$ whose $2n$ knots $s_0,t_0,\dots,s_{n-1},t_{n-1}$ are equally spaced on the unit circle $\{z:\ |z|=1\}$, but then the authors applied a variant of fast numerically stable FMM to compute HSS compressed approximation of this matrix and consequently to yield approximate solution of the original task in nearly linear arithmetic time. "HSS" and "FMM" are the acronyms for "Hierarchically Semiseparable" and "Fast Multipole Method", respectively. "Historically HSS representation is just a special case of the representations commonly exploited in the FMM literature" [CDGLP06]. We refer the reader to the papers [CGR98], [GR87], [DGR96], [BY13] and the bibliography therein on FMM and to [B10], [CDGLP06], [CGS07], [DV98], [GKK85], [T00], [X13], [XXG12], [XXCB], and the bibliography therein on HSS matrices and their link to FMM.

We analyze the fast numerically stable algorithms of [MRT05], [CGS07], [XXG12], and [XXCB], which treat the cited special Cauchy matrix and extend these algorithms to treat a quite general subclass of Cauchy matrices, which includes *CV* and *CV-like* matrices, obtained by FFT-based transforms from Vandermonde matrices and their transposes. The known approximation algorithms run in quadratic arithmetic time even for multiplication of these matrices by a vector, whereas we yield nearly linear arithmetic time both for that task and computing approximate solutions of linear systems of equations with these matrices where they are nonsingular and well conditioned. The solutions are immediately extended to the computations with Vandermonde matrices and to polynomial and rational evaluation and interpolation. In the cases of solving linear systems and interpolation, the power of our numerical algorithms is limited because only Vandermonde matrices of a narrow although important subclass are well conditioned (see [GI88]), and we prove a similar property for the CV matrices (see our Remark 31).

In Section 9.8 we employ another kind of transformations of matrix structures, which we call *functional*, to extend the power of the fast numerically stable algorithm of [DGR96], proposed for polynomial evaluation at a set of real knots. This enables us to accelerate a little further the approximation algorithms of [MRT05], [CGS07], [XXG12], and [XXCB] for Toeplitz linear systems. By means of other transformations of matrix structures we extend the approximation algorithms from CV matrices to Cauchy and Cauchy-like matrices with arbitrary sets of knots, but point out potential numerical limitations of these results. At the end of the paper we discuss some specific directions to the acceleration of our proposed approximation algorithms by logarithmic factor and to a more significant acceleration in the case of high precision computations. Further research could reveal new transformations of matrix structures with significant algorithmic applications.

Besides new demonstration of the power of the transformation techniques, our analysis of the approximation algorithms of [MRT05], [CGS07], [XXG12], and [XXCB] can be of some technical interest because instead of the special Cauchy matrix used in these papers we cover CV matrices, which have one of their two knot sets $\{s_0, \dots, s_{n-1}\}$ or $\{t_0, \dots, t_{n-1}\}$ equally spaced on the unit circle $\{z : \ |z| = 1\}$, whereas the remaining n knots are arbitrary. We still obtain a desired HSS representation by partitioning the knots according to the angles in their polar coordinates. Our study provides a new insight into the subject, and we prove that the admissible blocks of the $n \times n$ HSS ϵ-approximations of CV matrices have ranks of order $O(\log(n/\epsilon))$, which decrease to $O(\log(1/\epsilon))$ in the case of approximation of the special Cauchy matrix linked to Toeplitz inputs.

We organize our presentation as follows. After recalling some definitions and basic facts on general matrices and on four classes of structured matrices $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, and $\mathcal{C}$ in the next three sections, we cover in some detail the transformations of matrix structures among these classes in Section 5. That study is not used in the second part of our paper (Sections 6–8), where we approximate Cauchy matrices by HSS matrices. Namely we define HSS matrices in Section 6, estimate numerical ranks of Cauchy and Cauchy-like matrices of a large class in Section 7, and extend these estimates to compute the HSS type approximations of these matrices in Section 8. In Section 9 we combine the results of the two parts as well as some functional transformations of matrix structures to devise approximation algorithms for computations with various structured matrices. We conclude the paper with Section 10. For simplicity we assume square structured matrices throughout, but our study can be readily extended to the case of rectangular matrices (cf. [Pa]).

2 Some definitions and basic facts

Hereafter "flop" stands for "arithmetic operation with real or complex numbers"; the concepts "large", "small", "near", "close", "approximate", "ill conditioned" and "well conditioned" are quantified in the context. Next we recall and extend some basic definitions and facts on computations with general and structured matrices (cf. [GL96], [S98], [P01]).

2.1 General matrices

$M = (m_{i,j})_{i,j=1}^{m,n}$ is an $m \times n$ matrix, M^T and M^H are its transpose and Hermitian transpose, respectively. $M^{-T} = (M^T)^{-1} = (M^{-1})^T$. $(B_0 \mid \dots \mid B_{k-1})$ and $(B_0 \ \dots \ B_{k-1})$ denote a $1 \times k$ block matrix with the blocks $B_0, \dots, B_{k-1}$. $\Sigma = \mathrm{diag}(\Sigma_0, \dots, \Sigma_{k-1}) = \mathrm{diag}(\Sigma_j)_{j=0}^{k-1}$ is a $k \times k$ block diagonal matrix with the diagonal blocks $\Sigma_0, \dots, \Sigma_{k-1}$, possibly rectangular. $\mathbf{e}_1, \dots, \mathbf{e}_n$ are the n coordinate vectors of a dimension n. $\mathbf{s} = (s_j)_{j=0}^{n-1} = \sum_{i=0}^{n-1} s_i \mathbf{e}_i$. $D_\mathbf{s} = \mathrm{diag}(\mathbf{s}) = \mathrm{diag}(s_i)_{i=0}^{n-1}$. $I = I_n = (\mathbf{e}_1 \mid \dots \mid \mathbf{e}_n)$ and $J = J_n = (\mathbf{e}_n \mid \dots \mid \mathbf{e}_1)$ are the $n \times n$ identity and reflection matrices, respectively. $J = J^T = J^{-1}$.

Preprocessors. For three nonsingular matrices P, M, and N and a vector $\mathbf{b}$, it holds that

$$M^{-1} = N(PMN)^{-1}P, \quad PMN\mathbf{y} = P\mathbf{b}, \quad \mathbf{x} = N\mathbf{y}. \tag{1}$$

Generators. Given an $m \times n$ matrix M of a rank r and an integer $l \geq r$, we have a nonunique expression $M = FG^T$ for pairs (F, G) of matrices of sizes $m \times l$ and $n \times l$, respectively. We call such a pair (F, G) a *generator of length* l for the matrix M.

Norm, conditioning, orthogonality, numerical rank. $||M|| = ||M||_2$ is the spectral norm of an $n \times n$ matrix $M = (m_{i,j})_{i,j=0}^{n-1}$. We write $|M| = \max_{i,j=0}^{n-1} |m_{i,j}|$. $||M|| = ||M^H|| \leq \sqrt{mn}\ |M|$. For a fixed tolerance τ, the τ*-rank* of a matrix M is the minimum rank of matrices in its τ-neighborhood, $\{W :\ |W - M| \leq \tau\}$. The *numerical rank* of a matrix is its τ-rank for a small positive τ. A matrix is *ill conditioned* if its rank exceeds its numerical rank. A matrix M is *unitary* or *orthogonal* if $M^H M = I$ or $MM^H = I$. It is *quasiunitary* if cM is unitary for a nonzero constant c. A vector $\mathbf{u}$ is unitary if and only if $||\mathbf{u}|| = 1$, and if so, we call it a *unit vector*.

2.2 DFT and f-circulant matrices

Even for moderately large integers n the entries of an $n \times n$ Vandermonde matrix $V_\mathbf{s}$ vary in magnitude greatly unless $|s_i| \approx 1$ for all i. Next we cover the quasiunitary Vandermonde matrices Ω and Ω^H and the related class of f-circulant matrices (cf. [BP94, Section 3.4], [P01, Section 2.3 and 2.6]).

$\omega_n = \exp(\frac{2\pi}{n}\sqrt{-1})$ denotes a primitive nth root of 1. Its powers $1, \omega_n, \dots, \omega_n^{n-1}$ are equally spaced on the unit circle $\{z : |z| = 1\}$. $\Omega = \Omega_n = (\omega_n^{ij})_{i,j=0}^{n-1}$ denotes the $n \times n$ matrix of *DFT*, that is of the *discrete Fourier transform* at n points. It holds that $\Omega\Omega^H = nI$, and so Ω, Ω^H, and $\Omega^{-1} = \frac{1}{n}\Omega^H$ are quasiunitary matrices, whereas $\frac{1}{\sqrt{n}}\Omega$ and $\frac{1}{\sqrt{n}}\Omega^H$ are unitary matrices.

$Z_f = \begin{pmatrix} \mathbf{0}^T & f \\ I_{n-1} & \mathbf{0} \end{pmatrix}$ is the $n \times n$ matrix of f-circular shift for a scalar f,

$$JZ_fJ = Z_f^T, \quad JZ_f^TJ = Z_f \tag{2}$$

for any pairs of scalars e and f, and if $f \neq 0$, then

$$Z_f^{-1} = Z_{1/f}^T. \tag{3}$$

$Z_f(\mathbf{v}) = \sum_{i=0}^{n-1} v_i Z_f^i$ is an f-circulant matrix, defined by its first column $\mathbf{v} = (v_i)_{i=0}^{n-1}$ and a scalar $f \neq 0$ and called circulant for $f = 1$. It is a Toeplitz matrix and can be called a *DFT-based Toeplitz matrix* in view of the following results.

Theorem 1. *(See [CPW74].) It holds that* $Z_{f^n}(\mathbf{v}) = V_f^{-1} D(V_f \mathbf{v}) V_f$ *provided that* $f \neq 0$, $\Omega = (\omega_n^{ij})_{i,j=0}^{n-1}$ *is the* $n \times n$ *matrix of DFT,* $D(\mathbf{u}) = \mathrm{diag}(u_i)_{i=0}^{n-1}$ *for a vector* $\mathbf{u} = (u_i)_{i=0}^{n-1}$, *and* $V_f = \Omega\, \mathrm{diag}(f^i)_{i=0}^{n-1}$ *is the matrix of (6). In particular* $Z_1(\mathbf{v}) = \Omega^{-1} D(\Omega \mathbf{v}) \Omega$.

2.3 Cauchy and Vandermonde matrices

Recall the following properties of Cauchy and Vandermonde matrices (cf. [P01, Chapters 2 and 3]),

$$C_{\mathbf{s},\mathbf{t}} = -C_{\mathbf{t},\mathbf{s}}^T, \tag{4}$$

$$C_{\mathbf{s},\mathbf{t}} = \operatorname{diag}(t(s_i)^{-1})_{i=0}^{n-1} V_{\mathbf{s}} V_{\mathbf{t}}^{-1} \operatorname{diag}(t'(t_j))_{j=0}^{n-1} \tag{5}$$

where $\mathbf{s} = (s_i)_{i=1}^{n}$, $\mathbf{t} = (t_j)_{j=0}^{n-1}$, and $t(x) = \prod_{j=0}^{n-1}(x - t_j)$.

Theorem 2. $\det(V) = \prod_{i>k}(s_i - s_k)$ *and* $\det(C) = \prod_{i<j}(s_j - s_i)(t_i - t_j)/\prod_{i,j}(s_i - t_j)$

Corollary 3. *The matrices V and C of Table 1 are nonsingular where all $2n$ scalars $s_0,\dots,s_{n-1}$, $t_0,\dots,t_{n-1}$ are distinct.*

Theorem 4. *A row interchange preserves both Vandermonde and Cauchy structures. A column interchange preserves Cauchy structure.*

Equations (4) and (5) link together Cauchy and Vandermonde matrices and their transposes. Next we simplify these links. Write

$$V_f = ((f\omega_n^i)^j)_{i,j=0}^{n-1} = \Omega \operatorname{diag}(f^j)_{j=0}^{n-1}, \tag{6}$$

$$C_{\mathbf{s},f} = \left(\frac{1}{s_i - f\omega_n^j}\right)_{i,j=0}^{n-1},\ C_{e,\mathbf{t}} = \left(\frac{1}{e\omega_n^i - t_j}\right)_{i,j=0}^{n-1},\ C_{e,f} = \left(\frac{1}{e\omega_n^i - f\omega_n^j}\right)_{i,j=0}^{n-1}$$

and observe that $\Omega = V_1$, $\Omega^H = V_1^{-1}$, and *the matrices V_f are quasiunitary* where $|f| = 1$.

For $\mathbf{t} = (f\omega_n^j)_{j=0}^{n-1}$, it holds that $t(x) = x^n - f^n$, $t'(x) = nx^{n-1}$, $t(s_i) = s_i^n - f^n$, $t'(t_j) = nf^{n-1}\omega_n^{-j}$ for all j, and $nV_f^{-1} = \operatorname{diag}(f^{-i})_{i=0}^{n-1}\Omega^H$. Substitute these equations into (5) and obtain

$$C_{\mathbf{s},f} = \operatorname{diag}\left(\frac{f^{n-1}}{s_i^n - f^n}\right)_{i=0}^{n-1} V_{\mathbf{s}} \operatorname{diag}(f^{-j})_{j=0}^{n-1}\Omega^H \operatorname{diag}(\omega_n^{-j})_{j=0}^{n-1}, \tag{7}$$

$$C_{e,f} = \frac{f^{n-1}}{e^n - f^n}\Omega \operatorname{diag}((e/f)^i)_{i=0}^{n-1}\Omega^H \operatorname{diag}(\omega_n^{-j})_{j=0}^{n-1},$$

$$V_{\mathbf{s}} = \frac{f^{1-n}}{n}\operatorname{diag}\left(s_i^n - f^n\right)_{i=0}^{n-1} C_{\mathbf{s},f} \operatorname{diag}(\omega_n^j)_{j=0}^{n-1}\Omega \operatorname{diag}(f^j)_{j=0}^{n-1}, \tag{8}$$

$$V_{\mathbf{s}}^T = -\frac{f^{1-n}}{n}\operatorname{diag}(f^j)_{j=0}^{n-1}\Omega \operatorname{diag}(\omega_n^j)_{j=0}^{n-1} C_{f,\mathbf{s}} \operatorname{diag}(s_i^n - f^n)_{i=0}^{n-1}, \tag{9}$$

$$V_{\mathbf{s}}^{-1} = n \operatorname{diag}(f^{-j})_{j=0}^{n-1}\Omega^H \operatorname{diag}(\omega_n^{-j})_{j=0}^{n-1} C_{\mathbf{s},f}^{-1} \operatorname{diag}\left(\frac{f^{n-1}}{s_i^n - f^n}\right)_{i=0}^{n-1}, \tag{10}$$

$$V_{\mathbf{s}}^{-T} = -n \operatorname{diag}\left(\frac{f^{n-1}}{s_i^n - f^n}\right)_{i=0}^{n-1} C_{f,\mathbf{s}}^{-1} \operatorname{diag}(\omega_n^{-j})_{j=0}^{n-1}\Omega^H \operatorname{diag}(f^{-j})_{j=0}^{n-1}. \tag{11}$$

Definition 5. Hereafter we refer to the matrices V_f, $C_{\mathbf{s},f}$, $C_{e,\mathbf{t}}$, and $C_{e,f}$ for two scalars e and f as *FV, FC, CF*, and *FCF matrices*, respectively. We refer to the matrices $C_{\mathbf{s},f}$ and $C_{e,\mathbf{t}}$ as *CV matrices* and to the FV matrices V_f and the FCF matrices $C_{e,f}$ as the *DFT-based* matrices.

Similarly to the DFT matrix Ω, the DFT-based matrices have their basic sets of knots $\mathbb{S} = \{s_1,\dots,s_n\}$ and $\mathbb{T} = \{t_1,\dots,t_n\}$ equally spaced on the unit circle $\{z:\ |z| = 1\}$, and equation (7) links the CV matrices to Vandermonde matrices. In spite of all these links Cauchy and Vandermonde matrices also have very distinct features (cf. Remark 48).

Finally [P01, equation (3.4.1)] links a Vandermonde matrix and its transpose as follows,

$$V_{\mathbf{t}} J Z_f(\mathbf{w} + f\mathbf{e}_1) V_{\mathbf{t}}^T = \operatorname{diag}(t'(t_i)(f - t_i^n))_{i=0}^{n-1} \text{ for any scalar } f. \tag{12}$$

Here $t(x) = \prod_{j=0}^{n-1}(x - t_j)$ and $w(x) = t(x) - x^n$ are polynomials with the coefficient vectors $\mathbf{t}$ and $\mathbf{w}$, respectively (see Theorem 45 on their evaluation and Example 46 on their approximation).

2.4 The complexity of computations with DFT, Toeplitz, Hankel, Cauchy and Vandermonde matrices

We begin with the following observation.

Theorem 6. *If T is a Toeplitz matrix, then TJ and JT are Hankel matrices, whereas If H is a Hankel matrix, then HJ and JH are Toeplitz matrices.*

We also recall the following results (see, e.g., [BP94, Sections 1.2 and 3.4] on their proof and on the numerical stability of the supporting algorithms).

Theorem 7. *For any vector $\mathbf{v}$ of dimension n one can compute the vectors $\Omega\mathbf{v}$ and $\Omega^{-1}\mathbf{v}$ by using $O(n\log(n))$ flops. If n is a power of 2, then one can compute the vectors $\Omega\mathbf{v}$ and $\Omega^{-1}\mathbf{v}$ by applying* FFT*, that is by using $0.5n\log_2(n)$ and $0.5n\log_2(n)+n$ flops, respectively.*

Theorems 1, 6, and 7 combined with various techniques of matrix computations, imply the following results (cf. [P01, Chapter 2 and 3]).

Theorem 8. *$O(n\log^h(n))$ flops are sufficient to compute the product of an $n\times n$ matrix M and a vector $\mathbf{u}$ where $h=1$ if M is a Toeplitz or Hankel matrix and $h=2$ if M is a Vandermonde matrix, its transpose, or a Cauchy matrix. $O(n\log^2(n))$ flops are sufficient to compute the solution $\mathbf{x}$ of a nonsingular linear system of n equations $M\mathbf{x}=\mathbf{u}$ with any of such matrices M.*

The algorithms supporting this theorem are numerically stable where the matrix M is DFT-based (combine Theorems 1 and 7 and equation (6)) and where we multiply a Toeplitz or Hankel matrix M by a vector (embed an $n\times n$ Toeplitz matrix into $(2n-1)\times(2n-1)$ circulant matrix and then combine Theorems 1, 6, and 7). Otherwise the algorithms have numerical stability problems, and for numerical computations the users employ quadratic arithmetic time algorithms [BEGO08], [BF00], [KZ08], in spite of substantial research progress reported in the papers [PRT92], [PSLT93], [P95], [PZHY97], and particularly [DGR96], which applied a 1-dimensional adaptive FMM using Lagrange interpolation at Chebyshev's knots to prove the following result.

Theorem 9. *(Cf. [DGR96, Sections 3 and 4].) Assume a positive $\epsilon<1$, a unit vector $\mathbf{u}$, and an $n\times n$ Cauchy matrix $C_{\mathbf{s},\mathbf{t}}$ with real knots s_i and t_j. Then some numerically stable algorithms use $O(n\log(1/\epsilon))$ flops to approximate within the norm bound ϵ the product $C_{\mathbf{s},\mathbf{t}}\mathbf{u}$ and if the matrix is nonsingular, then also the solution $\mathbf{x}$ of the linear system $C_{\mathbf{s},\mathbf{t}}\mathbf{x}=\mathbf{u}$.*

3 The structures of Toeplitz, Hankel, Vandermonde and Cauchy types. Displacement ranks and generators

We generalize the four classes of matrices of Table 1 by employing the Sylvester displacements $AM-MB$ where the pair of *operator matrices* A and B is associated with a fixed matrix structure. (See [P01, Theorem 1.3.1] on a simple link to the Stein displacements $M-AMB$.) The rank and the generators of the displacement of a matrix M (for a fixed operator matrices A and B and tolerance τ) are said to be the *displacement rank* (denoted $d_{A,B}(M)$) and the *displacement generators*, of the matrix M, respectively (cf. [KKM79]), [P01], [BM01]).

Definition 10. If the displacement rank of a matrix is small (in context) for a pair of operator matrices associated with Toeplitz, Hankel, Vandermonde, transpose of Vandermonde or Cauchy matrices in Theorem 12 below, then the matrix is said to have the *structure of Toeplitz, Hankel, Vandermonde, transposed Vandermonde or Cauchy type*, respectively. Hereafter $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, $\mathcal{V}^T$, and $\mathcal{C}$ denote the five classes of these matrices (cf. Table 2). The classes $\mathcal{V}$, $\mathcal{V}^T$, and $\mathcal{C}$ consist of distinct subclasses $\mathcal{V}_{\mathbf{s}}$, $\mathcal{V}^T_{\mathbf{s}}$, and $\mathcal{C}_{\mathbf{s},\mathbf{t}}$ defined by the vectors $\mathbf{s}$ and $\mathbf{t}$ and the operator matrices $D_{\mathbf{s}}$ and $D_{\mathbf{t}}$, respectively, or equivalently by the bases $V_{\mathbf{s}}$ and $C_{\mathbf{s},\mathbf{t}}$ of these subclasses. To simplify the notation we will sometimes drop the subscripts $\mathbf{s}$ and $\mathbf{t}$ where they are not important or are defined by context.

Definition 11. (Cf. Definition 5.) Define the matrix classes $\mathcal{FV} = \cup_e \mathcal{V}_e$, $\mathcal{FC} = \cup_f \mathcal{C}_{\mathbf{s},f}$, $\mathcal{CF} = \cup_e \mathcal{C}_{e,\mathbf{t}}$, and $\mathcal{FCF} = \cup_{e,f} \mathcal{C}_{e,f}$ where the unions are over all complex scalars e and f. These matrix classes extend the classes of FV, FC, CF, and FCF matrices, respectively. We also define the classes $\mathcal{CV}$ (extending the CV matrices) and $\mathcal{V}^T\mathcal{F} = \cup_e \mathcal{V}_e^T$. We say that the above matrix classes consist of FV-like, V^TF-like, FC-like, CF-like, FCF-like, and CV-like matrices, which have structures of $\mathcal{FV}$-type, $\mathcal{V}^T\mathcal{F}$-type, $\mathcal{FC}$-type, $\mathcal{CF}$-type, $\mathcal{FCF}$-type, and $\mathcal{CV}$-type, respectively.

In our Theorems 12 and 14 we write (t), (h), (th), (v), (v^T), and (c) to indicate the matrix structures of Toeplitz, Hankel, Toeplitz or Hankel, Vandermonde, transposed Vandermonde, and Cauchy types, respectively. Recall the following well known results.

Theorem 12. Displacements of basic structured matrices.

(th) For a pair of scalars e and f and two matrices T (Toeplitz) and H (Hankel) of Table 1, the following displacements have ranks at most 2 (see some expressions for the shortest displacement generators in [P01, Section 4.2]),

$$Z_eT - TZ_f,\ Z_e^TT - TZ_f^T,\ Z_e^TH - HZ_f \text{ and } Z_eH - HZ_f^T.$$

(v) For a scalar e and a Vandermonde matrix V of Table 1 we have

$$VZ_e = D_{\mathbf{s}}V - (s_i^n - e)_{i=0}^{n-1}\mathbf{e}_n^T, \tag{13}$$

$$Z_e^TV^T = V^TD_{\mathbf{s}} - \mathbf{e}_n((s_i^n - e)_{i=0}^{n-1})^T, \tag{14}$$

and so the displacements $D_{\mathbf{s}}V - VZ_e$ and $Z_e^TV^T - V^TD_{\mathbf{s}}$ either vanish if $s_i^n = e$ for $i = 0, \ldots, n-1$ or have rank 1 otherwise.

(c) For two vectors $\mathbf{s} = (s_i)_{i=0}^{n-1}$ and $\mathbf{t} = (t_i)_{i=0}^{n-1}$, a Cauchy matrix C of Table 1, and the vector $\mathbf{e} = (1, \ldots, 1)^T$ of dimension n filled with ones, it holds that

$$D_{\mathbf{s}}C - CD_{\mathbf{t}} = \mathbf{e}\mathbf{e}^T, \quad \text{rank}(D_{\mathbf{s}}C - CD_{\mathbf{t}}) = 1. \tag{15}$$

Theorem 13. *For two scalars e and f and five matrices A, B, C, D, and M we have $d_{C,D}(M) - d_{A,B}(M) \le 1$ where either $A = C$, $B = Z_e$, $D = Z_f$ or $A = C$, $B = Z_e^T$, $D = Z_f^T$ and similarly where either $B = D$, $A = Z_e$, $C = Z_f$ or $B = D$, $A = Z_e^T$, $C = Z_f^T$.*

Proof. The matrix $Z_e - Z_f = (e - f)\mathbf{e}_1\mathbf{e}_n^T$ has rank at most 1 for any pair of scalars e and f. Therefore the matrices $(Z_eM - MB) - (Z_fM - MB) = Z_eM - Z_fM = (Z_e - Z_f)M$ and $(AM - MZ_e) - (AM - MZ_f) = -M(Z_e - Z_f)$ have ranks at most 1. □

The theorem implies that the classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, and $\mathcal{V}^T$ stay intact when we vary the scalars e and f, defining the operator matrices Z_e and Z_f.

Table 2 displays the pairs of operator matrices associated with the matrices of the seven classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}_{\mathbf{s}}$, $\mathcal{V}_{\mathbf{s}}^{-1}$, $\mathcal{V}_{\mathbf{s}}^T$, $\mathcal{V}_{\mathbf{s}}^{-T}$, and $\mathcal{C}_{\mathbf{s},\mathbf{t}}$. Five of these classes are employed in Theorems 12 and 14. $\mathcal{V}_{\mathbf{s}}^{-1}$ and $\mathcal{V}_{\mathbf{s}}^{-T}$ denote the classes of the inverses and the transposed inverses of the matrices of the class $\mathcal{V}_{\mathbf{s}}$, respectively. Equation (17) of the next section enables us to express their associated operator matrices through the ones for the classes $\mathcal{V}_{\mathbf{s}}$ and $\mathcal{V}_{\mathbf{s}}^T$.

The following theorem expresses the n^2 entries of an $n \times n$ matrix M through the $2dn$ entries of its displacement generator (F, G) defined under the operator matrices of Theorem 12 and Table 2. See some other expressions for various classes of structured matrices through their generators in [GO94], [P01, Sections 4.4 and 4.5], and [PW03].

Theorem 14. *Suppose $s_0, \ldots, s_{n-1}, t_0, \ldots, t_{n-1}$ are $2n$ distinct scalars, $\mathbf{s} = (s_k)_{k=0}^{n-1}$, $\mathbf{t} = (t_k)_{k=0}^{n-1}$, $V = (s_i^{k-1})_{i,k=0}^{n-1}$, $C = (\frac{1}{s_i - t_k})_{i,k=0}^{n-1}$, e and f are two distinct scalars, $\mathbf{f}_1, \ldots, \mathbf{f}_d, \mathbf{g}_1, \ldots, \mathbf{g}_d$ are $2d$ vectors of dimension n, $\mathbf{u}_1, \ldots, \mathbf{u}_n, \mathbf{v}_1, \ldots, \mathbf{v}_n$ are $2n$ vectors of dimension d, and F and G are $n \times d$ matrices such that* $F = \begin{pmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{pmatrix} = (\mathbf{f}_1 \mid \cdots \mid \mathbf{f}_d)$, $G = \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{pmatrix} = (\mathbf{g}_1 \mid \cdots \mid \mathbf{g}_d)$. *Then*

Table 2: Operator matrices for the seven classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}_{\mathbf{s}}$, $\mathcal{V}_{\mathbf{s}}^{-1}$, $\mathcal{V}_{\mathbf{s}}^{T}$, $\mathcal{V}_{\mathbf{s}}^{-T}$, and $\mathcal{C}_{\mathbf{s},\mathbf{t}}$

$\mathcal{T}$	$\mathcal{H}$	$\mathcal{V}_{\mathbf{s}}$	$\mathcal{V}_{\mathbf{s}}^{-1}$	$\mathcal{V}_{\mathbf{s}}^{T}$	$\mathcal{V}_{\mathbf{s}}^{-T}$	$\mathcal{C}_{\mathbf{s},\mathbf{t}}$
(Z_e, Z_f)	(Z_e^T, Z_f)	$(D_{\mathbf{s}}, Z_e)$	$(Z_e, D_{\mathbf{s}})$	$(Z_e^T, D_{\mathbf{s}})$	$(D_{\mathbf{s}}, Z_e^T)$	$(D_{\mathbf{s}}, D_{\mathbf{t}})$
(Z_e^T, Z_f^T)	(Z_e, Z_f^T)					

(t) $(e-f)M = \sum_{j=1}^{d} Z_e(\mathbf{f}_j)Z_f(J\mathbf{g}_j)$ *if* $Z_eM - MZ_f = FG^T$, $e \neq f$;

$(e-f)M = \sum_{j=1}^{d} Z_e(J\mathbf{f}_j)^T Z_f(\mathbf{g}_j)^T = J\sum_{j=1}^{d} Z_e(J\mathbf{f}_j)Z_f(\mathbf{g}_j)J$ *if* $Z_e^TM - MZ_f^T = FG^T$, $e \neq f$,

(h) $(e-f)M = \sum_{j=1}^{d} Z_e(\mathbf{f}_j)Z_f(\mathbf{g}_j)J$ *if* $Z_eM - MZ_f^T = FG^T$, $e \neq f$;

$(e-f)M = J\sum_{j=1}^{d} Z_e(J\mathbf{f}_j)Z_f(J\mathbf{g}_j)^T$ *if* $Z_e^TM - MZ_f = FG^T$, $e \neq f$,

(v) $M = \operatorname{diag}(\frac{1}{s_i^n - e})_{i=0}^{n-1} \sum_{j=1}^{d} \operatorname{diag}(\mathbf{f}_j)VZ_e(J\mathbf{g}_j)$ *if* $D_{\mathbf{s}}M - MZ_e = FG^T$ *and if* $s_i^n \neq e$ *for* $i = 0, \ldots, n-1$;

(v^T) $M = \operatorname{diag}(\frac{1}{e - s_i^n})_{i=0}^{n-1} \sum_{j=1}^{d} Z_e(J\mathbf{f}_j)^T V^T \operatorname{diag}(\mathbf{g}_j)$ *if* $Z_e^TM - MD_{\mathbf{s}} = FG^T$ *and if* $s_i^n \neq e$ *for* $i = 0, \ldots, n-1$;

(c) $M = \sum_{j=1}^{d} \operatorname{diag}(\mathbf{f}_j)C\operatorname{diag}(\mathbf{g}_j) = \left(\frac{\mathbf{u}_i^T\mathbf{v}_j}{s_i - t_j}\right)_{i,j=0}^{n-1}$ *if* $D_{\mathbf{s}}M - MD_{\mathbf{t}} = FG^T$.

Proof. Parts (t) and (h) are taken from [P01, Examples 4.4.2 and 4.4.5]. Part (c) is taken from [P01, Example 1.4.1]. To prove part (v), combine the equations $D_{\mathbf{s}}M - MZ_e = FG^T$ and $Z_eZ_{1/e}^T = I$ (cf. (3)) and deduce that $M - D_{\mathbf{s}}MZ_{1/e}^T = -F(Z_{1/e}G)^T$. Then obtain from [P01, Example 4.4.6 (part b)] that $M = e\operatorname{diag}(\frac{1}{s_i^n - e})_{i=0}^{n-1}\sum_{j=1}^{d}\operatorname{diag}(\mathbf{f}_j)VZ_{1/e}(Z_{1/e}\mathbf{g}_j)^T$. Substitute $eZ_{1/e}(Z_{1/e}\mathbf{g}_j) = Z_e(J\mathbf{g}_j)^T$ and obtain the claimed expression of part (v). Next transpose the equation $Z_e^TM - MD_{\mathbf{t}} = FG^T$ and yield $D_{\mathbf{s}}M^T - M^TZ_e = -GF^T$. From part (v) obtain $M^T = \operatorname{diag}(\frac{1}{e - s_i^n})_{i=0}^{n-1}\sum_{j=1}^{d}\operatorname{diag}(\mathbf{g}_j)VZ_e(J\mathbf{f}_j)$. Transpose this equation and arrive at part (v^T). □

By combining Theorems 8 and 14 we obtain the following results.

Theorem 15. *Given a vector* $\mathbf{v}$ *of a dimension* n *and a displacement generator of a length* d *for a matrix* M, *one can compute the product* $M\mathbf{v}$ *by using* $O(dn\log(n))$ *flops for an* $n \times n$ *matrix* M *in the classes* $\mathcal{T}$ *or* $\mathcal{H}$ *and by using* $O(dn\log^2(n))$ *flops for an* $n \times n$ *matrix* M *in* $\mathcal{V}$, $\mathcal{V}^T$, *or* $\mathcal{C}$.

Remark 16. By virtue of Theorem 14 the displacement operators $M \to AM - MB$ are nonsingular provided that $e \neq f$ in parts (t) and (h) and provided that $t_i^n \neq e$ for $i = 0, \ldots, n-1$ in parts (v) and (v^T). We can apply Theorem 13 to satisfy these assumptions.

Remark 17. (Cf. Theorem 4.) Parts (v) and (c) of Theorem 14 imply that a row interchange preserves the matrix structures of the Vandermonde and Cauchy types, whereas a column interchange preserves the matrix structures of the transposed Vandermonde and Cauchy types.

4 Matrix operations in terms of displacement generators

To accelerate pairwise multiplication and the inversion of structured matrices of large sizes we express them as well as the intermediate and final results of the computations through short displacement generators rather than the matrix entries. Such computations are possible by virtue of the following simple results from [P00] and [P01, Section 1.5], extending [P90].

Theorem 18. *Assume five matrices* A, B, C, M *and* N. *Then*

$$A(MN) - (MN)C = (AM - MB)N + M(BN - NC) \text{ and } \quad (16)$$

$$AM^{-1} - M^{-1}B = -M^{-1}(BM - MA)M^{-1} \quad (17)$$

provided that the matrix multiplications and inversion involved are well defined.

Table 3: Operator matrices for matrix product

P	M	N	PMN
C	A	B	C
A	B	D	D

Corollary 19. *For five matrices A, B, F, G, and M of sizes $m\times m$, $n\times n$, $m\times d$, $n\times d$, and $m\times m$, respectively, let us write $F = F_{A,B}(M)$, $G = G_{A,B}(M)$, and $d = d_{A,B}(M)$ if $AM - MB = FG^T$. Then under the assumptions of Theorem 18 we obtain the following equations,*

$$F_{A,B}(M^T) = -G_{B^T,A^T}(M^T),\ G_{A,B}(M^T) = F_{B^T,A^T}(M^T),$$
$$F_{A,C}(MN) = (F_{A,B}(M) \mid MF_{B,C}(N)),$$
$$G_{A,C}(MN) = (N^T G_{A,B}(M) \mid G_{B,C}(N)),$$
$$F_{A,B}(M^{-1}) = -M^{-1}G_{B,A}(M),\ G_{A,B}(M^{-1}) = M^{-T}F_{B,A}(M),$$

and so $d_{A,B}(M^T) = d_{B^T,A^T}(M)$, $d_{A,C}(MN) \le d_{A,B}(M) + d_{B,C}(N)$, $d_{A,B}(M^{-1}) = d_{B,A}(M)$.

The corollary and Theorem 14 together reduce the inversion of a nonsingular $n\times n$ matrix M given with a displacement generator of a length d to solving $2d$ linear systems of equations with this coefficient matrix M, rather than the n linear systems $M\mathbf{x}_i = \mathbf{e}_i$, $i = 1,\dots,n$.

Given short displacement generators for the matrices M and N, we can apply Corollary 19 and readily express short displacement generators for the matrices M^T and MN through the matrices M and N and their displacement generators, but the expressions for the displacement generator of the inverse M^{-1} involve the inverse itself.

5 Transformations of Displacement Matrix Structures

Equations (6), (7), (8)–(11) link Cauchy and Vandermonde matrix structures together, by means of multiplication by structured matrices. We are going to generalize this technique. We begin with recalling some simple links among Toeplitz, Hankel, and Vandermonde matrices. Then we will cover the approach comprehensively.

Theorem 20. *(i) JH and HJ are Toeplitz matrices if H is a Hankel matrix, and vice versa. (ii) $V^TV = (\sum_{k=0}^{m-1} s_k^{i+j})_{i,j=0}^{n-1}$ is a Hankel matrix for any $m\times n$ Vandermonde matrix $V = (s_i^j)_{i,j=0}^{m-1,n-1}$.*

5.1 Maps and multipliers

Recall that each of the five matrix classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, $\mathcal{V}^T$, and $\mathcal{C}$ consists of the matrices M whose displacement rank, rank$(AM - MB)$ is small (in context) for a pair of operator matrices (A, B) associated with this class and representing its structure. Theorem 18 shows the impact of elementary matrix operations on the associated operator matrices A and B. The operations of transposition and inversion change the associated pair (A, B) into $(-B^T, A^T)$ or $(-B, A)$. If the inputs of the operations are in any of the classes $\mathcal{T}$, $\mathcal{H}$, and $\mathcal{C}_{\mathbf{s},\mathbf{t}}$, then so are the outputs. Furthermore the transposition maps the classes $\mathcal{V}$ and $\mathcal{V}^T$ into one another, whereas inversion maps them into the classes $\mathcal{V}^{-1}$ and $\mathcal{V}^{-T}$, respectively. The impact of multiplication on matrix structure is quite different. By virtue of (16) and Table 3, the map $M \to PMN$ can define the transition from the associated pair of operator matrices (A, B) to any new pair (C, D) of our choice, that is we can transform the matrix structures of the five classes into each other at will. The following theorem and Table 4 specify such transforms of the structures given by the maps $M \to MN$, $N \to MN$, and $M \to PMN$ for appropriate multipliers P, M, and N.

Table 4: Mapping matrix structures by means of multiplication

$\mathcal{T}$	$\mathcal{H}$	$\mathcal{V}_{\mathbf{s}}$	$\mathcal{V}_{\mathbf{s}}^T$	$\mathcal{C}_{\mathbf{s},\mathbf{t}}$
$\mathcal{T}\mathcal{T}$, $\mathcal{V}_{\mathbf{s}}^T\mathcal{V}_{\mathbf{s}}^{-T}$	$\mathcal{T}\mathcal{H}$, $\mathcal{V}_{\mathbf{s}}^T\mathcal{V}_{\mathbf{s}}$	$\mathcal{V}_{\mathbf{s}}\mathcal{T}$, $\mathcal{C}_{\mathbf{s},\mathbf{t}}\mathcal{V}_{\mathbf{t}}$	$\mathcal{T}\mathcal{V}_{\mathbf{s}}^T$, $\mathcal{H}\mathcal{V}_{\mathbf{s}}^{-1}$	$\mathcal{V}_{\mathbf{s}}^{-T}\mathcal{V}_{\mathbf{t}}^T$, $\mathcal{V}_{\mathbf{s}}\mathcal{V}_{\mathbf{t}}^{-1}$, $\mathcal{C}_{\mathbf{s},\mathbf{q}}\mathcal{C}_{\mathbf{q},\mathbf{t}}$
$\mathcal{V}_{\mathbf{s}}^{-1}\mathcal{V}_{\mathbf{s}}$, $\mathcal{H}\mathcal{H}$	$\mathcal{H}\mathcal{T}$, $\mathcal{V}_{\mathbf{s}}^{-1}\mathcal{V}_{\mathbf{s}}^{-T}$	$\mathcal{V}_{\mathbf{s}}^{-T}\mathcal{H}$	$\mathcal{V}_{\mathbf{q}}^T\mathcal{C}_{\mathbf{q},\mathbf{s}}$	$\mathcal{V}_{\mathbf{s}}\mathcal{H}\mathcal{V}_{\mathbf{t}}^T$, $\mathcal{V}_{\mathbf{s}}^{-T}\mathcal{H}\mathcal{V}_{\mathbf{t}}^{-1}$

Theorem 21. *It holds that*

(i) $MN \in \mathcal{T}$ if the pair of matrices (M, N) is in any of the pairs of matrix classes $(\mathcal{T}, \mathcal{T})$, $(\mathcal{H}, \mathcal{H})$, $(\mathcal{V}_{\mathbf{s}}^{-1}, \mathcal{V}_{\mathbf{s}})$ and $(\mathcal{V}_{\mathbf{s}}^T, \mathcal{V}_{\mathbf{s}}^{-T})$,

(ii) $MN \in \mathcal{H}$ if the pair (M, N) is in any of the pairs $(\mathcal{T}, \mathcal{H})$, $(\mathcal{H}, \mathcal{T})$, $(\mathcal{V}_{\mathbf{s}}^{-1}, \mathcal{V}_{\mathbf{s}}^{-T})$ and $(\mathcal{V}_{\mathbf{s}}^T, \mathcal{V}_{\mathbf{s}})$,

(iii) $MN \in \mathcal{V}_{\mathbf{s}}$ if the pair (M, N) is in any of the pairs $(\mathcal{V}_{\mathbf{s}}, \mathcal{T})$, $(\mathcal{V}_{\mathbf{s}}^{-T}, \mathcal{H})$, and $(\mathcal{C}_{\mathbf{s},\mathbf{t}}, \mathcal{V}_{\mathbf{t}})$,

(iv) $MN \in \mathcal{V}_{\mathbf{s}}^T$ if the pair (M, N) is in any of the pairs $(\mathcal{T}, \mathcal{V}_{\mathbf{s}}^T)$, $(\mathcal{H}, \mathcal{V}_{\mathbf{s}}^{-1})$ and $(\mathcal{V}_{\mathbf{q}}^T, \mathcal{C}_{\mathbf{q},\mathbf{s}})$,

(v) $MN \in \mathcal{C}_{\mathbf{s},\mathbf{t}}$ if the pair (M, N) is in any of the pairs $(\mathcal{C}_{\mathbf{s},\mathbf{q}}, \mathcal{C}_{\mathbf{q},\mathbf{t}})$, $(\mathcal{V}_{\mathbf{s}}^{-T}, \mathcal{V}_{\mathbf{s}}^T)$ and $(\mathcal{V}_{\mathbf{s}}, \mathcal{V}_{\mathbf{s}}^{-1})$,

(vii) $PMN \in \mathcal{C}_{\mathbf{s},\mathbf{t}}$ if the triple (M, N, P) is in any of the triples $(\mathcal{V}_{\mathbf{s}}, \mathcal{H}, \mathcal{V}_{\mathbf{t}}^T)$ and $(\mathcal{V}_{\mathbf{s}}^{-T}, \mathcal{H}, \mathcal{V}_{\mathbf{t}}^{-1})$.

The maps of Theorem 21 and Table 4 hold for any choice of the multipliers P and N from the indicated classes. To simplify the computation of displacement generators for the products MN, MP and MNP, we can choose the multipliers J, $V_{\mathbf{r}}$, $V_{\mathbf{r}}^T$, $V_{\mathbf{r}}^{-1}$, $V_{\mathbf{r}}^{-T}$, and $C_{\mathbf{p},\mathbf{r}}$, all having displacement rank 1, to represent the classes $\mathcal{H}$, $\mathcal{V}_{\mathbf{r}}$, $\mathcal{V}_{\mathbf{r}}^T$, $\mathcal{V}_{\mathbf{r}}^{-1}$, $\mathcal{V}_{\mathbf{r}}^{-T}$, and $\mathcal{C}_{\mathbf{p},\mathbf{r}}$, respectively, where $\mathbf{p}$ and $\mathbf{r}$ can stand for $\mathbf{q}$, $\mathbf{s}$, and $\mathbf{t}$. Hereafter we call this choice of multipliers *canonical.* We call them *canonical and DFT-based* if up to the factor J they are also DFT-based, that is if $\mathbf{p}$ and $\mathbf{r}$ are of the form $f(\omega_n^i)_{i=0}^{n-1}$. These multipliers are quasiunitary where $|f| = 1$. By combining Corollary 19 and Theorem 21 we obtain the following result.

Corollary 22. *Suppose a displacement generator of a length d is given for an $n \times n$ matrix M of any of the classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, $\mathcal{V}^T$, and $\mathcal{C}$. Then $O(dn \log^2(n))$ flops are sufficient to compute a displacement generator of a length at most $d + 2$ for the matrix PMN of any other of these classes where P and M are from the set of canonical multipliers complemented by the identity matrix. The flop bound decreases to $O(dn \log(n))$ where the canonical multipliers are DFT-based.*

One can simplify the inversion of structured matrices M of some important classes and the solution of linear systems $M\mathbf{x} = \mathbf{u}$ by employing preprocessing $M \to PMN$ with appropriate structured multipliers P and N.

5.2 The impact on displacements

Theorems 23 and 24 of this subsection imply that in the canonical maps of Theorem 21 the displacement ranks grow by at most 2 but possibly less than that. Our *constructive proofs of these theorems also specify the multipliers P and N and the displacement generators* for the products PMN involved into the maps of Theorem 21. In the maps supporting parts (a)–(e) of the following theorem we set $P = I$ or $N = I$, thus omitting one of the multipliers. The theorem implicitly covers the maps where the matrices M or PMN belong to the classes $\mathcal{V}^T$, $\mathcal{V}^{-1}$, or $\mathcal{V}^{-T}$, because we can generate these maps by transposing or inverting the maps for $M \in \mathcal{V}$ and $PMN \in \mathcal{V}$.

Theorem 23. *Suppose a displacement generator of a length d is given for a structured matrix M of any of the four classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, and $\mathcal{C}$. Then one can obtain a displacement generator of a length at most $d + 2$ for a matrix PMN belonging to any other of these classes by selecting appropriate canonical multipliers P and N among the matrices I (from the class $\mathcal{T}$), J (from the class $\mathcal{H}$), Vandermonde matrices V and their transposes V^T. Namely, if we assume canonical multipliers P and N, then we can compute a displacement generator of the matrix PMN having a length at most d where the map $M \to PMN$ is between the matrices M and PMN in the classes $\mathcal{H}$ and $\mathcal{T}$. This length bound grows to at most $d + 2$ where M is in the class $\mathcal{T}$ or $\mathcal{H}$, whereas $PMN \in \mathcal{C}$ or vice versa, where $M \in \mathcal{C}$ and PMN is in the class $\mathcal{T}$ or $\mathcal{H}$. Displacement generators have lengths at most $d + 1$ in the maps $M \to PMN$ for all other transitions among the classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$ and $\mathcal{C}$.*

Proof. We specify some maps $M \to MNP$ that support the claims of the theorem. One can vary and combine these maps as well as the other maps of Theorem 21 and Table 4.

(a) $\mathcal{T} \to \mathcal{H}$, $PMN = JM$. Assume a matrix $M \in \mathcal{T}$, a pair of distinct scalars e and f, and a pair of $n \times d$ matrices $F = F_{Z_e,Z_f}(M)$ and $G = G_{Z_e,Z_f}(M)$ for $d = d_{Z_e,Z_f}(M)$ satisfying the displacement equation $Z_eM - MZ_f = FG^T$ (cf. Theorem 14). Pre-multiply this equation by the matrix J to obtain $JZ_eM - (JM)Z_f = JFG^T$. Rewrite the term $JZ_eM = JZ_eJJM$ as Z_e^TJM by observing that $JZ_eJ = Z_e^T$ (cf. (2) for $f = e$). Obtain $Z_e^T(JM) - (JM)Z_f = JFG^T$. Consequently $F_{Z_e^T,Z_f}(JM) = JF$, $G_{Z_e^T,Z_f}(JM) = G$, $d_{Z_e^T,Z_f}(JM) = d_{Z_e,Z_f}(M)$, and $JM \in \mathcal{H}$.

(b) $\mathcal{T} \to \mathcal{V}$, $PMN = VM$. Keep the assumptions of part (a) and fix n scalars $s_0, \dots, s_{n-1}$. Pre-multiply the displacement equation $Z_eM - MZ_f = FG^T$ by the Vandermonde matrix $V = (s_i^j)_{i,j=0}^{n-1}$ to obtain $VZ_eM - (VM)Z_f = VFG^T$. Write $\mathbf{s} = (s_i)_{i=0}^{n-1}$ and substitute equation (13) to yield $D_{\mathbf{s}}(VM) - (VM)Z_f = VFG^T + (s_i^n - e)_{i=0}^{n-1}\mathbf{e}_n^TM = F_{VM}G_{VM}^T$ for $F_{VM} = (VF \mid (s_i^n - e)_{i=0}^{n-1})$ and $G_{VM}^T = \begin{pmatrix} G^T \\ \mathbf{e}_n^TM \end{pmatrix}$. So $d_{D_{\mathbf{s}},Z_f}(VM) \le d_{Z_e,Z_f}(M) + 1$ and $VM \in \mathcal{V}_{\mathbf{s}}$.

(c) $\mathcal{H} \to \mathcal{T}$, $PMN = MJ$. Assume a matrix $M \in \mathcal{H}$, a pair of scalars e and f, and a pair of $n \times d$ matrices F and G for $d = d_{Z_e,Z_f^T}(M)$ satisfying the displacement equation $Z_eM - MZ_f^T = FG^T$ (cf. Theorem 14). Post-multiply it by the matrix J to obtain $Z_e(MJ) - MZ_f^TJ = FG^TJ$. Express the term $MZ_f^TJ = MJJZ_f^TJ$ as MJZ_f (cf. (2)) to obtain $Z_e(MJ) - (MJ)Z_f = FG^TJ = F(JG)^T$ and consequently $F_{Z_e,Z_f}(JM) = F$, $G_{Z_e,Z_f}(JM) = JG$, $d_{Z_e,Z_f}(MJ) = d_{Z_e,Z_f^T}(M)$ and $MJ \in \mathcal{T}$.

(d) $\mathcal{H} \to \mathcal{V}$. Compose the maps of parts (c) and (b).

(e) $\mathcal{V} \to \mathcal{H}$, $PMN = V^TM$. Assume $n + 2$ scalars $e, f, s_0, \dots, s_{n-1}$, a matrix $M \in \mathcal{V}$, and its displacement generator given by $n \times d$ matrices F and G such that $D_{\mathbf{s}}M - MZ_f = FG^T$ (cf. (13)). Pre-multiply this equation by the transposed Vandermonde matrix $V^T = (s_j^i)_{i,j=0}^{n-1}$ to obtain $V^TD_{\mathbf{s}}M - (V^TM)Z_f = V^TFG^T$ for $\mathbf{s} = (s_i)_{i=0}^{n-1}$. Apply equation (14) to express the matrix $V^TD_{\mathbf{s}}$ and obtain $Z_e^T(V^TM) - (V^TM)Z_f = V^TFG^T + \mathbf{e}_n((s_i^n - e)_{i=0}^{n-1})^TM = F_{V^TM}G_{V^TM}^T$ for $F_{V^TM} = (V^TF \mid \mathbf{e}_n)$ and $G_{V^TM}^T = \begin{pmatrix} G^T \\ ((s_i^n - e)_{i=0}^{n-1})^TM \end{pmatrix}$. So $d_{Z_e^T,Z_f}(V^TM) \le d_{D_{\mathbf{s}},Z_f}(M) + 1$ and $V^TM \in \mathcal{H}$.

(f) $\mathcal{V} \to \mathcal{T}$. Compose the maps of parts (e) and (c).

(g) $\mathcal{V} \to \mathcal{C}$, $PMN = MJV^T$. Assume $2n + 1$ scalars e, $s_0, \dots, s_{n-1}, t_0, \dots, t_{n-1}$, a matrix $M \in \mathcal{V}$, and its displacement generator given by $n \times d$ matrices F and G. Post-multiply the equation $D_{\mathbf{s}}M - MZ_e = FG^T$ (cf. (13)) by the matrix JV^T where $V^T = (t_j^i)_{i,j=0}^{n-1}$ is the transposed Vandermonde matrix, substitute $Z_eJ = JZ_e^T$, and obtain $D_{\mathbf{s}}(MJV^T) - MJZ_e^TV^T = FG^TJV^T$ for $\mathbf{s} = (s_i)_{i=0}^{n-1}$. Apply equation (14) to express the matrix $Z_e^TV^T$ and obtain $D_{\mathbf{s}}(MJV^T) - (MJV^T)D_{\mathbf{t}} = FG^TJV^T - MJ\mathbf{e}_n((t_i^n - e)_{i=0}^{n-1})^T = F_{MJV^T}G_{MJV^T}^T$ where $F_{MJV^T} = (F \mid MJ\mathbf{e}_n)$ and $G_{MJV^T}^T = \begin{pmatrix} G^TJV^T \\ ((e - t_i^n)_{i=0}^{n-1})^T \end{pmatrix}$. So $d_{D_{\mathbf{s}},D_{\mathbf{t}}}(MJV^T) \le d_{D_{\mathbf{s}},Z_e}(M) + 1$ and $MJV^T \in \mathcal{C}_{\mathbf{s},\mathbf{t}}$.

We can alternatively write $PMN = MV^{-1}$ for $V = V_{\mathbf{t}}$. (The matrix V can be readily inverted where it is FFT-based, that is where $V = V_f$.) Post-multiply the equation $D_{\mathbf{s}}M - MZ_e = FG^T$ (cf. (13)) by the matrix $V^{-1} = V_{\mathbf{t}}^{-1}$, for $\mathbf{t} = (t_i)_{i=0}^{n-1}$, to obtain $D_{\mathbf{s}}MV^{-1} - MZ_eV^{-1} = FG^TV^{-1}$. Pre- and post-multiply by V^{-1} equation (13) for $\mathbf{s}$ replaced by $\mathbf{t}$ and obtain $Z_eV^{-1} = V^{-1}D_{\mathbf{t}} - V^{-1}(t_i^n - e)_{i=0}^{n-1}\mathbf{e}_n^TV^{-1}$. Substitute this expression for Z_eV^{-1} into the above equation and obtain $D_{\mathbf{s}}(MV^{-1}) - (MV^{-1})D_{\mathbf{t}} = FG^TV^{-1} - V^{-1}(e - t_i^n)_{i=0}^{n-1}\mathbf{e}_n^TV^{-1} = F_{MV^{-1}}G_{MV^{-1}}^T$ for $F_{MV^{-1}} = (F \mid V^{-1}(e - t_i^n)_{i=0}^{n-1})$ and $G_{V^{-1}M}^T = \begin{pmatrix} G^TV^{-1} \\ \mathbf{e}_n^TV^{-1} \end{pmatrix}$. So $d_{D_{\mathbf{s}},D_{\mathbf{t}}}(VM) \le d_{D_{\mathbf{s}},Z_e}(M) + 1$ and $MV^{-1} \in \mathcal{C}_{\mathbf{s},\mathbf{t}}$.

(h) $\mathcal{C} \to \mathcal{V}$, $PMN = MV$. Assume $2n + 1$ scalars e, $s_0, \dots, s_{n-1}, t_0, \dots, t_{n-1}$, a matrix $M \in \mathcal{C}$, and its displacement generator given by $n \times d$ matrices F and G such that $D_{\mathbf{s}}M - MD_{\mathbf{t}} = FG^T$ for $\mathbf{s} = (s_i)_{i=0}^{n-1}$ and $\mathbf{t} = (t_i)_{i=0}^{n-1}$ (cf. (15)). Post-multiply this equation by the Vandermonde matrix $V = (t_i^{j-1})_{i,j=0}^{n-1}$ to obtain $D_{\mathbf{s}}(MV) - MD_{\mathbf{t}}V = FG^TV$. Express the matrix $D_{\mathbf{t}}V$ from matrix equation (13) and obtain $D_{\mathbf{s}}(MV) - (MV)Z_e = FG^TV + M(t_i^n - e)_{i=0}^{n-1}\mathbf{e}_n^T = F_{MV}G_{MV}^T$ where $F_{MV} = (F \mid M(t_i^n - e)_{i=0}^{n-1})$ and $G_{MV}^T = \begin{pmatrix} G^TV \\ \mathbf{e}_n^T \end{pmatrix}$. So $d_{D_{\mathbf{s}},Z_e}(MV) \le d_{D_{\mathbf{s}},D_{\mathbf{t}}}(M) + 1$ and $MV \in \mathcal{V}_{\mathbf{s}}$.

(i) $\mathcal{C} \to \mathcal{T}$. Compose the maps of parts (h) and (f).
(j) $\mathcal{C} \to \mathcal{H}$. Compose the maps of parts (h) and (e).
(k) $\mathcal{T} \to \mathcal{C}$. Compose the maps of parts (b) and (g).
(i) $\mathcal{H} \to \mathcal{C}$. Compose the maps of parts (d) and (g). □

Multiplications by a Cauchy matrix keeps a matrix in any of the classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$ and $\mathcal{C}$, but changes a diagonal operator matrix. Next we specify the impact on the displacement.

Theorem 24. *Assume $2n$ distinct scalars $s_0, \dots, s_{n-1}, t_0, \dots, t_{n-1}$, defining two vectors $\mathbf{s} = (s_i)_{i=0}^{n-1}$ and $\mathbf{t} = (t_j)_{j=0}^{n-1}$ and a nonsingular Cauchy matrix $C = C_{\mathbf{s},\mathbf{t}} = (\frac{1}{s_i - t_j})_{i,j=0}^{n-1} n$ (cf. part (i) of Theorem 2). Then for any pair of operator matrices A and B we have*
(i) $d_{A,D_\mathbf{t}}(MC) \le d_{A,D_\mathbf{s}}(M) + 1$ and
(ii) $d_{D_\mathbf{s},B}(CM) \le d_{D_\mathbf{t},B}(M) + 1$.

Proof. (i) We have $d_{A,D_\mathbf{s}}(M) = \text{rank}(AM - MD_\mathbf{s}) = \text{rank}(AMC - MD_\mathbf{s}C)$. Furthermore $AMC - MCD_\mathbf{t} = AMC - MD_\mathbf{s}C + MD_\mathbf{s}C - MCD_\mathbf{t} = (AM - MD_\mathbf{s})C + M(D_\mathbf{s}C - CD_\mathbf{t})$. Substitute equation (15) and deduce that $AMC - MCD_\mathbf{t} = (AM - MD_\mathbf{s})C + M\mathbf{e}\mathbf{e}^T$. Therefore $d_{A,D_\mathbf{t}}(MC) = \text{rank}(AMC - MCD_\mathbf{t}) \le \text{rank}((AM - MD_\mathbf{s})C) + 1 = \text{rank}(AM - MD_\mathbf{s}) + 1 = d_{A,D_\mathbf{s}}(M) + 1$.

(ii) We have $D_\mathbf{s}CM - CMB = D_\mathbf{s}CM - CD_\mathbf{t}M + CD_\mathbf{t}M - CMB = (D_\mathbf{s}C - CD_\mathbf{t})M + C(D_\mathbf{t}M - MB)$. Substitute equation (15) and deduce that $D_\mathbf{s}CM - CMB = C(D_\mathbf{t}M - MB) + \mathbf{e}\mathbf{e}^T M$. Therefore $d_{D_\mathbf{s},B}(CM) = \text{rank}(D_\mathbf{s}CM - CMB) \le \text{rank}(C(D_\mathbf{t}M - MB)) + 1 = \text{rank}(D_\mathbf{t}M - MB) + 1 = d_{D_\mathbf{t},B}(M) + 1$. □

5.3 Canonical and DFT-based transformations of the matrices of the classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$ and $\mathcal{V}^T$ into CV-like matrices

By combining equations (13), (14), and (16) one can deduce that multiplication by a Vandermonde multiplier $V = (s_i^{j-1})_{i,j=0}^{n-1}$ or by its transpose increases the length of a displacement generator by at most 1, but equations (13) and (14) imply that such multiplication does not increase the length at all where $s_i^n = e$ for $i = 0, \dots, n-1$ and for a scalar e, employed in the operator matrices Z_e and Z_e^T of the Vandermonde displacement map. This suggests choosing the vectors $\mathbf{s} = (e\omega_n^{i-1})_{i=0}^{n-1}$ and $\mathbf{t} = (f\omega_n^{i-1})_{i=0}^{n-1}$ and employing the DFT-based multipliers V_e and V_f of (6), in particular in our maps supporting part (g) of Theorem 23. Then the output matrices of the class $\mathcal{CV}$ would have the same displacement ranks as the input matrices M. Furthermore the inverse of the matrix $V = V_\mathbf{t}$, employed in our second map supporting part (g), would turn into DFT-based matrix V_f, and we could invert it and multiply it by a vector by using $O(n \log(n))$ flops (cf. Theorem 8). We deduce the following results by reexamining the proof of Theorem 23 and applying transposition.

Theorem 25. *The canonical DFT-based multipliers from the proof of Theorem 23 for the basic vectors $\mathbf{s} = (e\omega_n^i)_{i=0}^{n-1}$ and $\mathbf{t} = (f\omega_n^i)_{i=0}^{n-1}$ and for some appropriate complex scalars e and f support the following transformations of matrix classes (in both directions), $\mathcal{T} \leftrightarrow \mathcal{FV} \leftrightarrow \mathcal{FCF}$, $\mathcal{H} \leftrightarrow \mathcal{FV} \leftrightarrow \mathcal{FCF}$, $\mathcal{V} \leftrightarrow \mathcal{CF}$, $\mathcal{V}^T \leftrightarrow \mathcal{FC}$, and $\mathcal{V} \cup \mathcal{V}^T \leftrightarrow \mathcal{CV}$. The multipliers are quasiunitary (and thus the transformations are numericaly stable) where $|e| = |f| = 1$.*

By combining our second map in the proof of part (g) of Theorem 23 with our map from its part (b) and choosing $\mathbf{t} = (f\omega_n^i)_{i=0}^{n-1}$, so that the $2n$ knots $s_0, t_0, \dots, s_{n-1}, t_{n-1}$ are equally spaced on the unit circle $\{z : \ |z| = 1\}$, we can obtain canonical DFT-based transforms $\mathcal{T} \to \mathcal{C} = \Omega\mathcal{T}\,\text{diag}(f^i)_{i=0}^{n-1}\Omega^H$, which are quasiunitary where $|f| = 1$. For $f = \omega_{2n}$ they turn into the celebrated map employed in the papers [H95], [GKO95], [G98], [MRT05], [R06], [CGS07], [XXG12]. The following theorem shows the implied map of the displacement generators (see the proofs of parts (b) and (g) of Theorem 23, using the second map supporting part (g)).

Theorem 26. *Suppose $Z_1M - MZ_{-1} = FG^T$ for an $n \times n$ matrix M and $n \times d$ matrices F and G and write $P = \Omega$, $N = D_0^H\Omega^H$, $C = PMN$, $D_0 = \text{diag}(\omega_{2n}^i)_{i=0}^{n-1}$, and $D = D_0^2 = \text{diag}(\omega_n^i)_{i=0}^{n-1}$. Then $DC - \omega_{2n}CD = F_CG_C^T$ for $F_C = \Omega F$ and $G_C = \Omega D_0 G$.*

The theorem and the supporting canonical DFT-based map $\mathcal{T} \rightarrow \mathcal{C}$ are a special case of Theorem 23 and its transforms of matrix structures, extending [P90]. In his letter of 1991, reproduced in [P11, Appendix C], G. Heinig has acknowledged studying the paper [P90], although in [H95] he deduced Theorem 26 from Theorem 1 rather than supplying more general results based on Theorem 23.

6 HSS matrices

In the next four sections we study HSS matrices, employ them to approximate Cauchy matrices, and combine these results with the displacement and functional transformations of matrix structures to devise more efficient algorithms.

Definition 27. As in Section 1, "*HSS*" stands for "hierarchically semiseparable". An $n \times n$ matrix is (l, u)-HSS if its diagonal blocks consist of $O((l + u)n)$ entries, if l is the maximum rank of all its subdiagonal blocks, and if u is the maximum rank of all its superdiagonal blocks, that is blocks of all sizes lying strictly below or strictly above the block diagonal, respectively.

HSS matrices extend the class of banded matrices and their inverses, and similar extensions are known under the names of matrices with a low Hankel rank, quasiseparable, weakly, recursively or sequentially semiseparable matrices, and rank structured matrices. See [B10], [CDGLP06], [CGS07], [DV98], [EGH13a], [EGH13b], [GKK85], [T00], [VVGM05], [VVM07], [VVM08], [X13], [XXG12], [XXCB], and the bibliography therein on the long history of the study of these matrix classes and see [B10], [BY13], [DGR96], [CGR98], [GR87], [LRT79], [P93], [PR93], and the bibliography therein on the related subjects of FMM, Matrix Compression, and Nested Dissection algorithms.

One can readily express the n^2 entries of an (l, u)-HSS matrix of size $n \times n$ via $O((l + u)n)$ parameters of a generalized generator and can multiply this matrix by a vector by using $O((l+u)n)$ flops. If the matrix is nonsingular, then its inverse is also an (l, u)-HSS matrix, and $O((l+u)^3 n)$ flops are sufficient to compute a generalized generator expressing it via $O((l + u)n)$ parameters. Having computed such a generator, one can solve a linear system with this matrix by using $O((l + u)n)$ additional flops. See [DV98], [EG02], [MRT05], [CGS07], [XXG12], [XXCB], [Pa], and the references therein on supporting algorithms and their efficient implementation. Our next goal is the design of fast approximation algorithms for CV matrices by means of their approximation by slightly generalized HSS matrices, to which we extend fast HSS algorithms.

7 Low-rank approximation of certain Cauchy matrices

Definition 28. (See [CGS07, page 1254].) A pair of complex points s and t is (θ, c)-*separated* for $0 \le \theta < 1$ and a complex point c (a center) if $|\frac{t-c}{s-c}| \le \theta$. Two sets of complex numbers $\mathbb{S}$ and $\mathbb{T}$ are (θ, c)-*separated* from one another if every pair of elements $s \in \mathbb{S}$ and $t \in \mathbb{T}$ is (θ, c)-separated from one another for the same pair (θ, c).

Theorem 29. *(Cf. [MRT05], [CGS07, Section 2.2].) Suppose $C = (\frac{1}{s_i - t_j})_{i,j=0}^{n-1}$ is a Cauchy matrix defined by two sets of parameters $\mathbb{S} = \{s_0, \ldots, s_{n-1}\}$ and $\mathbb{T} = \{t_0, \ldots, t_{n-1}\}$. Suppose these sets are* (θ, c)-separated *from one another for $0 < \theta < 1$ and a center c and write*

$$\delta = \delta_{c,\mathbb{S}} = \min_{i=0}^{n-1} |s_i - c|. \tag{18}$$

Then for every positive integer k it is sufficient to use $2kn + 2n - 2$ flops to compute two matrices $F = (1/(s_i - c)^{h+1})_{i,h=0}^{n-1,k}$, $G^T = ((t_j - c)^h)_{j,h=0}^{n,k}$ that support the representation of the matrix C as $C = \widehat{C} + E$ where $\widehat{C} = FG^T$, $\operatorname{rank}(\widehat{C}) \le k + 1$, $|E| \le \frac{\theta^k}{(1-\theta)\delta}$ for all pairs $\{i, j\}$.

Remark 30. We can replace $\delta = \delta_{c,\mathbb{S}} = \min_{i=0}^{n-1} |s_i - c|$ by $\delta = \delta_{c,\mathbb{T}} = \min_{j=0}^{n-1} |t_j - c|$ because $C_{\mathbf{s},\mathbf{t}}^T = -C_{\mathbf{t},\mathbf{s}}$ (cf. (4)).

Remark 31. Unless the values $1-\theta>0$ and δ of (18) are small, the upper bound of Theorem 29 on the norm $|E|$ is small already for moderately large integers k. Then Theorem 29 implies an upper bound $k+1$ on the numerical rank of the large subclass of Cauchy matrices $C=(\frac{1}{s_i-t_j})_{i,j=0}^{n-1}$ whose parameter sets $\mathbb{S}=\{s_0,\dots,s_{n-1}\}$ and $\mathbb{T}=\{t_0,\dots,t_{n-1}\}$ are (θ,c)-separated from one another for an appropriate center c. Even if this property holds just for a subset of the set $\mathbb{S}$ that defines an $l\times n$ Cauchy submatrix where $l>k+1$, then this submatrix and consequently the matrix C as well are ill conditioned. In particular since all knots $t_0,\dots,t_{n-1}$ of a CV matrix lie on the unit circle $\{z:\ |z|=1\}$, they are $(\theta,0)$-separated (with θ not close to 1) from every knot s_i not lying close to this circle, and so a CV matrix is ill conditioned unless all but at most l of its knots s_i lie on or near this circle.

8 Local low-rank approximation of CV matrices

Theorem 29 defines a low-rank approximation of a CV matrix where its two knot sets are separated by a *global center* c. Generally a CV matrix has no such center, but next we show that it always has a set of *local centers* that support approximation by generalized HSS matrices. We begin with a simple lemma that expresses the distances between the two points of the unit circle $\{z:\ |z|=1\}$ and from its point to a sector.

Lemma 32. *Suppose* $0\le\phi<\phi'<\phi''\le 2\pi$, $\phi'-\phi\le\phi-\phi''+2\pi$, $\tau=\exp(\phi\sqrt{-1})$, $\tau'=\exp(\phi'\sqrt{-1})$, *and* $\tau''=\exp(\phi''\sqrt{-1})$ *and let* $\Gamma(\phi',\phi'')=\{r\exp(\mu\sqrt{-1}):\ r\ge 0,\ 0\le\phi'\le\mu<\phi''\le 2\pi\}$ *denote the semi-open sector on the complex plane bounded by two rays from the origin passing through the points* τ' *and* τ''. *Then (i)* $|\tau'-\tau|=2\sin((\phi'-\phi)/2)$ *and (ii) the distance from the point* τ *to the sector* $\Gamma(\phi',\phi'')$ *is equal to* $\sin(\phi'-\phi)$.

Theorem 33. *Assume sufficiently large positive integers* k, h *and* n, *a complex scalar* e, *and a CV matrix* $C=C_{\mathbf{s},e}=(\frac{1}{s_i-t_j})_{i,j=0}^{n-1}$ *such that* $t_j=e\omega_n^j$ *for* $j=0,\dots,n-1$, $kh=n$, *the integers* k *and* h *are not small, and* $e\ne 0$. *Then there is an* $n\times n$ *permutation matrix* P *such that* PC *is a block vector* $PC=(C_0,\dots,C_{k-1})$ *where a* basic block column C_p *has size* $n\times h$. *Furthermore consider the first row of the matrix* C *adjacent to its last row (as if they were glued together). Then every basic block column* C_p *can be partitioned into an* $\widehat{n}_p\times h$ extended diagonal block $\widehat{\Sigma}_p$ *and an* $(\widehat{n}-n_p)\times h$ admissible block $\widehat{N}_p$ *(see Remark 34), such that the blocks* $\widehat{\Sigma}_0,\dots,\widehat{\Sigma}_{k-1}$ *have* $3hn$ *entries overall, whereas every admissible block* $N_p=(\frac{1}{s_i-t_j})_{i\in\widehat{\mathbb{S}}_q,j\in\mathbb{T}_q}$ *is associated with a pair of knot sets* $\widehat{\mathbb{S}}_q\subseteq\mathbb{S}$ *and* $\mathbb{T}_q\subset\mathbb{T}$ *that are* (θ,c_p)*-separated from one another for a center* $c_p=\exp(\psi_p\sqrt{-1})$ *and* $\theta=2\sin(\mu)/\sin(\nu)\approx\tilde{\theta}=2\mu/\nu$ *where* $2\mu=\max\{|c_p-\omega_n^{ph}|,\ |c_p-\omega_n^{(p+1)h}|\}$, $\nu=\min\{|2(p-1)\pi k-\psi_p|,\ |2(p+2)\pi k-\psi_p|\}$, *and* ψ_p *is any number satisfying* $2p\pi/k\le\psi_p<2(p+1)\pi/k$. *In particular* $\mu=0.5\pi/k$, $\nu=3\pi/k$, *and* $\tilde{\theta}=1/3$ *provided* c_p *is the midpoint of the arc* $\mathbb{A}_p$ *of the unit circle* $\{z:\ |z|=1\}$ *with the endpoints* ω_n^{ph} *and* $\omega_n^{(p+1)h}$, *that is provided* $\psi_p=(2p+1)\pi/k$, *whereas* $\mu\le 0.75\pi/k$, $\nu\ge 2.5\pi/k$, *and* $\tilde{\theta}=3/5$ *provided* c_p *is a point on the arc* $\mathbb{A}'_p$ *with the end points* $\omega_{2n}^{(2p+0.5)h}$ *and* $\omega_{2n}^{(2p+1.5)h}$, *that is provided* $(2p+0.5)\pi/k\le\psi_p\le(2p+1.5)\pi/k$.

Proof. With no loss of generality assume that $e=1$ because the claimed properties are readily extended from the matrix $C_{\mathbf{s},1}$ to the matrix $\frac{1}{e}C_{\mathbf{s},1}=C_{e\mathbf{s},e}$. Represent the knots $s_0,\dots,s_{n-1}$ of the set $\mathbb{S}$ in polar coordinates, $s_i=r_i\exp(2\pi\phi_i\sqrt{-1})$ where $r_i\ge 0$, $0\le\phi_i<2\pi$, $\phi_i=0$ if $r_i=0$, and $i=0,1,\dots,n-1$. Re-enumerate all values ϕ_i to have them in nondecreasing order and to have $\phi_0^{(\mathrm{new})}=\min_{i=0}^{n}\phi_i$ and let P denote the permutation matrix that defines this re-enumeration. To simplify our notation assume that already the original enumeration has these properties, that is $P=I$. Let $\mathbb{S}_p$ and $\mathbb{T}_p$ denote the two subsets of the sets $\mathbb{S}$ and $\mathbb{T}$, respectively, that lie in the semi-open sector of the complex plane $\Gamma_p=\{z=r\exp(\psi\sqrt{-1}):\ r\ge 0,\ 2\pi p/k\le\psi<2\pi(p+1)/k\}$, bounded by the pair of the rays from the origin to the points ω_n^{ph} and $\omega_n^{(p+1)h}$. Define the block partition $C=(C_{p,q})_{p,q=0}^{k-1}$ with the blocks $C_{p,q}=(\frac{1}{s_i-t_j})_{i\in\mathbb{S}_p,j\in\mathbb{T}_q}$ for $p,q=0,\dots,k-1$. Then declare the first block row of the matrix $(C_{p,q})_{p,q=0}^{k-1}$ adjacent to its last row, that is declare the blocks $C_{0,q}$

and $C_{k-1,q}$ pairwise adjacent for every q (as if the two rows were glued together), and partition every block column $C_{.,q} = (C_{p,q})_{q=0}^{k-1}$ into the extended diagonal block $\widehat{\Sigma}_q = (\frac{1}{s_i-t_j})_{i\in\mathbb{S}'_q, j\in\mathbb{T}_q}$ and the admissible block $\widehat{N}_q = (\frac{1}{s_i-t_j})_{i\in\widehat{\mathbb{S}}_q, j\in\mathbb{T}_q}$, where $\widehat{\mathbb{S}}_q = \mathbb{S} - \mathbb{S}'_q$, $\mathbb{S}'_q = \mathbb{S}_{q-1 \mod k} \cup \mathbb{S}_q \cup \mathbb{S}_{q+1 \mod k}$, $q = 0, \dots, k-1$. Clearly the k diagonal blocks $\Sigma_p = C_{p,p}$ of sizes $n_p \times h$ for $p = 0, \dots, k-1$ have $h\sum_{p=0}^{k-1} n_p = hn$ entries overall, and this overall number is tripled in the extension to the blocks $\widehat{\Sigma}_p$ because $\widehat{n}_p = n_{p-1 \mod k} + n_q + n_{p+1 \mod k}$, and so $h\sum_{p=0}^{k-1} \widehat{n}_p = 3h\sum_{p=0}^{k-1} n_p = 3hn$. Furthermore Lemma 32 implies that the sets $\widehat{\mathbb{S}}_q$ and $\mathbb{T}_q$ defining the admissible block $\widehat{N}_q$ are (θ, c_q)-separated from one another for θ and c_p defined in the theorem. □

Remark 34. Every block $\widehat{\Sigma}_p$ and $\widehat{N}_p$ is made up of a pair blocks of the matrix CP that are either adjacent to one another or become adjacent if we declare that the first row of the matrix is adjacent to its last row.

Remark 35. Even if k does not divide n we can still partition the unit circle by k equally spaced points, then partition the complex plane into k sectors accordingly, represent the matrix C as a $k \times k$ block matrix $(C_{p,q})_{p,q=0}^{k-1}$, and define the basic block columns and the diagonal, extended diagonal, and admissible blocks. The only change in the claims and proofs is that we would allow the number of columns of the latter blocks to vary slightly, by at most 1, as q varies. The techniques of the adaptive FMM [CGR98] enable us to handle the case of any distribution of the knots s_i on a circle as well as on a line or smooth curve of a bounded length. See Section 9 on further extensions.

Remark 36. (Recursive merging.) Our analysis and results hold for any positive integer k. We recursively apply them by partitioning the unit circle by arcs whose lengths increase at every recursive step. More precisely, at every recursive step we merge a pair of the adjacent arcs of the current finer partition of the circle into a single arc of the new coarser partition. Then we redefine the diagonal, extended diagonal and admissible blocks. This *recursive merging* dramatically enhances the power of Theorem 33 for supporting fast CV algorithms.

Remark 37. (Bounding the distance from the centers to the knot set $\mathbb{S}$.) There are exactly n elements $s_0, \dots, s_{n-1}$ in the set $\mathbb{S}$. Therefore for every p we can choose a center c_p on the arc $\mathbb{A}'_p$ at the distance at least $2\sin(\pi/(8kn))$ from this set and thus obtain the bound

$$\delta \geq \delta_- = 2\sin(\pi/(8kn) \tag{19}$$

for δ of (18) (cf. part (i) of Lemma 32), where $\delta_- \approx \pi/(4kn)$ for large n. Alternatively we can choose the centers c_p at the midpoints of the arcs $\mathbb{A}_p$, and rotate both arcs $\mathbb{A}_p$ and centers c_p for $p = 0, \dots, k-1$ by a fixed angle on the unit circle $\{z : |z| = 1\}$. For a proper choice of the angle, part (i) of Lemma 32 ensures that $\delta \geq 2\sin(\pi/(4kn))$ where $2\sin(\pi/(4kn) \approx \pi/(2kn) \geq 0.25\pi/n^2$ for large integers n. This would have decreased bound (19) by a factor of 2, but to support the application of Theorem 33 throughout the merging process, we use about $2n$ centers c_p overall. This makes the same impact on the value δ_- as halving the length of the arcs $\mathbb{A}_p$ and brings us back to bound (19).

Combine Theorem 29 with bound (19) and obtain the following result.

Corollary 38. *At the kth stage of recursive merging, for $1 < k < n$, every admissible block $\widehat{N}_q$ of the matrix PC of Theorem 33 can be ϵ-approximated by a matrix of rank ρ provided that $\epsilon = \frac{4\theta^\rho}{(1-\theta)\delta\pi}$ for δ of Theorem 29. For a constant θ, $0 < \theta < 1$ this holds where $\rho = O(\log(\frac{1}{\delta\epsilon}))$ and consequently, by virtue of (19), where*

$$\rho = O(\log(n/\epsilon)). \tag{20}$$

Remark 39. Equation (4) implies that Theorem 33 and consequently Corollary 38 and Theorem 43 of the next section can be immediately extended to the case where $C = (\frac{1}{s_i-t_j})_{i,j=0}^{n-1}$, $s_i = e\omega^i$ for all i, and the choice of the knots t_j is unrestricted. One can apply the FMM techniques of [GR87] and [CGR98] toward relaxing our restrictions on the knot sets $\mathbb{S}$ or $\mathbb{T}$ of a Cauchy matrix. They would guarantee numerical stability, unlike our alternative techniques in Section 9.5.

Remark 40. The lower bound $\delta_- \approx \pi/(2km)$ on δ of Remark 37 is overly pessimistic for many dispositions of the knots s_i on the complex plane. For example, δ_- is a constant where the value $|s_i|$ is close to 1 for no i, whereas $\delta_- \geq \pi/m$ where $s_i = \omega_m^i$, $i = 0, \ldots, m-1$. Furthermore typically at most a small fraction of all differences $c_{j,q} - s_i$ has absolute values close to the bound δ_-. For constant δ_- bound (20) on the ϵ-rank decreases to $\rho = O(\log(1/\epsilon))$, which is the case in [DGR96, Section 3], where all knots s_i and t_j are real.

9 Fast approximate computations with structured matrices and extensions

9.1 Definitions and auxiliary results

We need some additional definitions and basic results. $\alpha(M)$ and $\beta(M)$ denote the arithmetic cost of computing the vectors $M\mathbf{u}$ and $M^{-1}\mathbf{u}$, respectively, maximized over all unit vectors $\mathbf{u}$ and minimized over all algorithms, and we write $\alpha_\epsilon(M) = \min_{|E|\leq\epsilon} \alpha(M+E)$ and $\beta_\epsilon(M) = \min_{|E|\leq\epsilon} \beta(M+E)$ for a fixed small positive ϵ. The straightforward algorithm supports the following bound.

Theorem 41. *$\alpha(M) \leq 2(m+n)\rho - \rho - m$ for an $m \times n$ matrix M given with its generating pair of a length ρ.*

Theorem 42. *(See [S98, Corollary 1.4.19] for $P = -M^{-1}E$.) Suppose M and $M+E$ are two nonsingular matrices of the same size and $||M^{-1}E|| = \theta < 1$. Then $||I - (M+E)^{-1}M|| \leq \frac{\theta}{1-\theta}$ and $||(M+E)^{-1} - M^{-1}|| \leq \frac{\theta}{1-\theta}||M^{-1}||$. In particular $||(M+E)^{-1} - M^{-1}|| \leq 1.5\theta||M^{-1}||$ if $\theta \leq 1/3$.*

9.2 Fast approximate computations with CV matrices

Unlike the case of HSS matrices Corollary 38 bounds numerical rank only for off-diagonal blocks defined column-wise, but not row-wise. We can still devise fast algorithms for such generalized HSS matrices approximating a CV matrix C because these column-wise bounds hold throughout the process of recursive merging. Our bounds on the cost of approximate solution of linear systems of equations actually require that the associated merging processes of this and the previous sections for an HSS approximation involve no singular or ill conditioned auxiliary matrices [Pa], and in particular the input matrix should be nonsingular and well conditioned. In view of Remark 31 this is a serious restriction, which is extended to our algorithms for solving structured linear systems of equations in the next subsections.

Theorem 43. *Assume an $n \times n$ CV matrix C, a positive $\epsilon < n$, and ρ of (20). Then $\alpha_\epsilon(C) = O(n\rho\log(n)))$ and $\beta_\epsilon(C) = O(n\rho^2\log(n))$, and so $\alpha_\epsilon(C) = O(n\log^2(n)))$ and $\beta_\epsilon(C) = O(n\log^3(n))$ where $\rho = O(\log(n))$, whereas $\alpha_\epsilon(C) = O(n\log(1/\epsilon)\log(n)))$ and $\beta_\epsilon(C) = O(n\log^2(1/\epsilon)\log(n))$ where $\rho = O(\log(1/\epsilon))$.*

Proof. Apply recursive merging to the matrix C. At its jth stage, $j = 0, \ldots, l-1$, for $l \leq \lceil \log_2 n \rceil$, compute a permutation matrix $P^{(j)}$ and an ϵ-approximation $C_\epsilon^{(j)}$ of the matrix $P^{(j)}C$ where all admissible blocks have ranks at most ρ for ρ and ϵ invariant at all stages of the merging process and satisfying equation (20) (cf. Corollary 38). Clearly $\alpha_\epsilon(C) \leq \alpha(C_\epsilon^{(j)})$ for all j. Let $C' = C_\epsilon^{(l-1)}$ denote the matrix entering the last merging stage, $C'_\epsilon = \begin{pmatrix} \widehat{\Sigma}'_0 & \widehat{N}_1 \\ \widehat{\Sigma}''_0 & \widehat{\Sigma}'_1 \\ \widehat{N}_0 & \widehat{\Sigma}''_1 \end{pmatrix}$ where $\widehat{\Sigma}_p = \begin{pmatrix} \widehat{\Sigma}'_p \\ \widehat{\Sigma}''_p \end{pmatrix}$ for $p = 0, 1$ denote the two extended diagonal blocks and where $\mathrm{rank}(\widehat{N}_p) \leq \rho$ for $p = 0, 1$. It follows that $\alpha_\epsilon(C) \leq \alpha(C'_\epsilon) \leq \sum_{p=0}^{1}(\alpha(\widehat{\Sigma}_p) + \alpha(\widehat{N}_p)) + n$. Apply Theorem 41 and obtain that $\sum_{p=0}^{1} \alpha(\widehat{N}_p) \leq 2n\rho$. Recursively apply this argument to estimate $\alpha(\widehat{\Sigma}_p)$ for $p = 0, 1$ and obtain $\alpha(C') \leq n + 4n\rho l + \sum_{j=0}^{l-1} \alpha(\widehat{\Sigma}_j)$ where $\widehat{\Sigma}_j$ denotes the matrix made up of the extended diagonal blocks at the jth merging, $j = 0, \ldots, l-1$ and having at most $3nh$ entries for every j. Choose $h = O(\rho)$ and obtain that $\sum_{j=0}^{l-1} \alpha(\widehat{\Sigma}_j) = O(nl\rho)$ and consequently $\alpha_\epsilon(C) \le \alpha(C') = O(n\rho\log(n))$. The claimed bound on $\beta_\epsilon(C)$ is supported by the algorithms of [CGS07] and [XXG12]. The algorithms have been proposed for HSS matrices, approximating the special CV matrix $C_{1,\omega_{2n}}$, but close examination in [Pa] shows that they support the claimed cost bound for CV matrices. □

9.3 Extension to Vandermonde matrices and their transposes

Theorem 44. *Suppose we are given a vector $\mathbf{s} = (s_i)_{i=0}^{n-1}$ defining an $n \times n$ Vandermonde matrix $V = V_\mathbf{s}$. Write $s_+ = \max_{i=0}^{n-1} |s_i|$ and $\bar\epsilon = (s_+ + 1)\epsilon$ where $\log(1/\epsilon) = O(\log n)$. Then $\alpha_{\bar\epsilon}(V) + \alpha_{\bar\epsilon}(V^T) = O(n\rho\log(n))$ and $\beta_{\bar\epsilon}(V) + \beta_{\bar\epsilon}(V^T) = O(n\rho^2\log(n))$ for ρ of (20).*

Proof. Equations (8)–(11) reduce the computations with the matrices $V_\mathbf{s}$ and $V_\mathbf{s}^T$ to the same computations with a Cauchy matrix $C_{\mathbf{s},f}$, which is a CV matrix for any f such that $|f| = 1$. This enables us to extend Theorem 43 to the matrices $V_\mathbf{s}$ and $V_\mathbf{s}^T$ except that we must adjust the approximation bound ϵ of that theorem. Let us show that it is sufficient to change it to $\bar\epsilon$. The matrices $\mathrm{diag}(\omega^{-j})_{j=0}^{n-1}$, $\mathrm{diag}(f^{-j})_{j=0}^{n-1}$, and $\frac{1}{\sqrt{n}}\Omega = (\sqrt{n}\Omega^H)^{-1}$ and their inverses are unitary, and so multiplication by them makes no impact on the output error norms. Multiplication by the matrix $\mathrm{diag}(s_i^n - f^n)_{i=0}^{n-1}$ can increase the value $\log_2(1/\epsilon)$ by at most $\log_2(s_+^n + 1)$, whereas multiplication by its inverse can increase this value by at most $\log_2(\Delta)$ for $\Delta = 1/\max_{\{f:\ |f|=1\}} \min_{i=0}^{n-1} |s_i^n - f^n|$. We can ensure that $\Delta \le 2n$ by choosing a proper value f, and so $\log_2(\Delta) \le 1 + \log_2(n)$. Such an increase makes no impact on the asymptotic bounds of Theorem 43. □

9.4 Extension to the classes of $\mathcal{CV}$, $\mathcal{V}$ and $\mathcal{V}^T$

Assume a Cauchy-like matrix M of the class $\mathcal{C}$ represented with its displacement generator (F, G) of a length d. Part (c) of Theorem 14 enables us to reduce the approximation of M to approximation of its basic matrix C. We immediately obtain that

$$\alpha_{\epsilon'}(M) \le 2d + d\alpha_\epsilon(C) \text{ for } \epsilon' \le \sum_{j=1}^{d} |\mathbf{f}_j|\ |\mathbf{g}_j|\ \epsilon \le d\ |F|\ |G|\ \epsilon. \tag{21}$$

Combine this estimate with Theorem 43 provided that M is a CV-like matrix and obtain

$$\alpha_{\epsilon'}(M) = O(dn\rho\log(n)) \tag{22}$$

for ρ, ϵ' and ϵ' satisfying (20) and (21). To estimate $\beta_{\epsilon'}(M)$ note that the ranks of the matrices of local approximation increase by at most a factor of d in the transition to the matrix M from its basic matrix C in the expression of part (c) of Theorem 14, whereas the approximation norm bounds are defined according to the expressions for ϵ and ϵ' of (21). Theorem 42 extends the latter error norm bounds to the case of the solution of linear systems with these matrices as follows,

$$\epsilon'' = \epsilon''(\epsilon', M) = O(n\epsilon'||M^{-1}||), \tag{23}$$

and we obtain the following expression,

$$\beta_{\epsilon''}(M) = O(d^2n\rho^2\log(n)) \tag{24}$$

where ρ, ϵ', and ϵ'' satisfy (20), (21) and (23). Furthermore we reduce the approximation of the matrices of the classes $\mathcal{V}$ and $\mathcal{V}^T$ to the approximation of CV matrices by applying at first parts (v) and (v^T) of Theorem 14 and then the algorithms supporting Theorem 44.

9.5 Extension to the case of arbitrary knots

Next we transform matrix structure based on Theorem 24 to extend our approximation algorithms to computations with Cauchy and Cauchy-like matrices of the class $\mathcal{C}$ with any set of knots, and we also estimate the impact of the transformations on the approximation errors.

Suppose that $\mathbf{s}$, $\mathbf{t}$ and $\mathbf{u}$ denote three vectors of dimension n and that an $n\times n$ Cauchy-like matrix $M \in \mathcal{C}_{\mathbf{s},\mathbf{t}}$ is given with a displacement generator (F, G) of a length d. Transform matrix structures to reduce the solution of a linear system $M\mathbf{x} = \mathbf{u}$ to some computations with CV-like matrices and multiplication of the matrix M by the vector $\mathbf{e} = (1, \dots, 1)^T$. Fix a scalar e, $|e| = 1$, write $P = MC_{\mathbf{t},e}$ and $\mathbf{x} = C_{\mathbf{t},e}\mathbf{y}$, and note that $P\mathbf{y} =\mathbf{u}$, where $P \in \mathcal{C}_{\mathbf{s},e}$ is a CV-like matrix with a displacement generator (F_P, G_P) of a length at most $d + 1$, $F_P = (F \mid M\mathbf{e})$ and $G_P = (C^T_{\mathbf{t},e}G \mid \mathbf{e})$. By applying these techniques to the matrix $M^T \in \mathcal{C}_{\mathbf{t},\mathbf{s}}$ we can alternatively reduce the linear system $M\mathbf{x} = \mathbf{u}$ to the computation of the products $M^T\mathbf{e}$ and to some computations with CV-like matrices. Likewise part (c) of Theorem 14 reduces the approximation of the vector $\mathbf{x} = M\mathbf{u}$ to the approximation of the d vectors $C_{\mathbf{s},\mathbf{t}}\mathbf{v}_i$ for $\mathbf{v}_i = \operatorname{diag}(\mathbf{g}_i)_{i=1}^d\mathbf{u}$, $\mathbf{g}_i = G\mathbf{e}_i$, and $i = 1, \dots, d$, and to $O(dn)$ additional flops, provided that the matrix $M \in \mathcal{C}_{\mathbf{s},\mathbf{t}}$ is given with its displacement generator (F, G) of a length d. We can compute a displacement generator of a length at most 2 for the matrix $C' = C_{\mathbf{s},\mathbf{t}}C_{\mathbf{t},e}$, of the class $\mathcal{C}_{\mathbf{s},e}$, and then reduce the computation of the vector $\mathbf{x}$ to multiplication of the $\mathcal{CV}$ matrix C' by the vector $\mathbf{z}$ satisfying the CV linear system of equations $C_{\mathbf{t},e}\mathbf{z} = \mathbf{u}$. In all cases we reduce the original tasks to computations with CV matrices and readily verify that multiplication by the auxiliary CV matrix C increases the approximation error norm of the output by at most a factor of $||C||\ ||C^{-1}||$. As we showed in Remark 31 this upper bound is large unless most of the knots of the CV matrix C have absolute values near 1, and so one may benefit from alternative direct applications of the FMM to Cauchy matrices.

9.6 Fast approximate computations with polynomials and rational functions

Together with our equations (7), (8), and (10), the following results link polynomial and rational interpolation and multipoint evaluation to each other, multiplication of Vandermonde and Cauchy matrices by a vector, and the solution of Vandermonde and Cauchy linear systems of equations (cf. [P01, Chapter 3]). By using this link we can extend our results on Vandermonde and Cauchy matrices to polynomial and rational evaluation and interpolation, respectively.

Theorem 45. *(i) Let $p(x) = \sum_{i=0}^{n-1} p_i x^i$, $\mathbf{p} = (p_i)_{i=0}^{n-1}$, $\mathbf{s} = (s_i)_{i=0}^{n-1}$, and $\mathbf{v} = (v_i)_{i=0}^{n-1}$. Then the equations $p(s_i) = v_i$ hold for $i = 0, 1, \dots, n-1$ if and only if $V_{\mathbf{s}}\mathbf{p} = \mathbf{v}$. (ii) For a rational function $v(x) = \sum_{j=0}^{n-1} \frac{u_j}{x-t_j}$ with n distinct poles $t_0, \dots, t_{n-1}$ and for n distinct scalars $s_0, \dots, s_{n-1}$, write $\mathbf{s} = (s_i)_{i=0}^{n-1}$, $\mathbf{t} = (t_j)_{j=0}^{n-1}$, $\mathbf{u} = (u_j)_{j=0}^{n-1}$, $\mathbf{v} = (v_i)_{i=0}^{n-1}$. Then the equations $v_i = v(s_i)$ hold for $i = 0, \dots, n-1$ if and only if $C_{\mathbf{s},\mathbf{t}}\mathbf{u} = \mathbf{v}$. (iii) The equation $t(x) = \prod_{j=0}^{n-1}(x - t_j) = x^n + w(x)$, for n distinct knots $t_0, \dots, t_{n-1}$, is equivalent to the linear systems of n equations $w(t_j) = -t_j^n$ or $n+1$ equations $t(0) = (-1)^n \prod_{j=0}^{n-1} t_j$, $t(t_j) = 0$ for $j = 0, \dots, n-1$ in both cases, that is to polynomial interpolation for the vectors $(-t_j)_{j=0}^{n-1}$ and $(t(0), 0, \dots, 0)^T$, respectively (cf. Example 46).*

9.7 Functional transformations of matrix structures and computations with generalized Cauchy matrices

Our algorithms of this section fall into the general framework of the FMM, and we can incorporate their modifications and extensions known in the FMM literature, such as application of Lagrange interpolation instead of Taylor expansion (cf. [DGR96], [B10]) and the extension to generalized input classes. Toward the latter extension, consider Cauchy matrices $C_{\mathbf{s},\mathbf{t}}$ as a discrete representation of the function $\frac{1}{s-t}$, transform this function into various other functions of the variable $s - t$ such as $a + \frac{b}{s-t-c}$, $\frac{1}{(s-t)^2}$, and $\ln(s - t)$, and arrive at various *generalized Cauchy matrices* such as $\big(a + \frac{b}{s_i-t_j-c}\big)_{i,j=0}^{n-1}$, $\big(\frac{1}{(s_i-t_j)^2}\big)_{i,j=0}^{n-1}$, and $\big(\ln(s_j - t_j)\big)_{i,j=0}^{n-1}$. We can readily extend to these matrices the FMM/HSS algorithms and complexity estimates (cf. [GR87], [DGR96]). In particular observe that $\frac{1}{(s-t)^2} = \frac{1}{(s-c)^2}\frac{1}{(1-q)^2}$, $\frac{az+b}{z-h} = a + \frac{b+ah}{z-h}$, and $\ln(s - t) = \ln(s - c) + \ln(1 - q)$ for $q = \frac{t-c}{s-c}$. Let us sketch an application to polynomial computations.

Example 46. Represent the polynomial $t(x) = \prod_{j=0}^{n-1}(x-t_j)$ of Section 2.3 and part (c) of Theorem 45 as $\exp(\sum_{j=0}^{n-1}\ln(x-t_j))$ and approximate its values $t(s_i) = \exp(\sum_{j=0}^{n-1}\ln(s_i-t_j))$ at the nth roots of unity $s_i = \omega^i$ by using the FMM/HSS techniques. Now apply IDFT to the computed approximations to the values $v(\omega^i) = t(\omega^i) - 1$ of the polynomial $v(x) = t(x) - x^n$ of a degree at most $n-1$ to approximate the coefficients of the polynomials $v(x)$ and consequently $t(x)$.

9.8 Extensions of Theorem 9 and approximate solution of Toeplitz and Hankel linear systems of equations

If all knots s_i and t_j of a Cauchy matrix $C_{\mathbf{s},\mathbf{t}}$ lie on the real line $\{z:\ \Im(z)=0\}$, then Theorem 9 decreases the bounds of Theorem 43 on $\alpha_\epsilon(C)$ and $\beta_\epsilon(C)$ by factors of $\log(\rho)$ and $\log^2(\rho)$, respectively, and decreases also the ϵ-rank ρ of the off-diagonal blocks from order of $\log(n/\epsilon)$ to $O(\log(1/\epsilon))$, bounding it just in terms of the error tolerance ϵ. Next we apply functional transformations of Cauchy matrices to extend Theorem 9 to cover the cases where the $2n$ knots can lie on any line or circle on the complex plane.

Theorem 47. *Theorem 9 holds for a Cauchy matrix $C_{\mathbf{s},\mathbf{t}}$ where all knots s_i and t_j lie on any line on the complex plane.*

Proof. Begin with the following observations where $a \neq 0$ and c are two complex constants,

$$C_{\mathbf{s},\mathbf{t}} = C_{\mathbf{s}',\mathbf{t}'} \text{ where } s'_i = s_i - c,\ t'_j = t_j - c \text{ for all } i \text{ and } j, \tag{25}$$

$$C_{\mathbf{s},\mathbf{t}} = aC_{\mathbf{s}',\mathbf{t}'} \text{ where } a \neq 0,\ s'_i = s_i/a,\ t'_j = t_j/a \text{ for all } i \text{ and } j, \tag{26}$$

Now suppose all knots s_i and t_j lie on a line obtained by rotating the real line by an angle ϕ followed by the shift by a complex c. Define the new knots $s'_i = (s_i - c)/a$ and $t'_j = (t_j - c)/a$ for $a = \exp(\phi\sqrt{-1})$ and all i and j. They lie on the real line. Apply Theorem 9 to the matrix $C_{\mathbf{s}',\mathbf{t}'}$, and apply equations (25) and (26) to extend the resulting approximations to the matrix $C_{\mathbf{s},\mathbf{t}}$. $\square$

Remark 48. Equations (25) and (26) show low impact of shift and scaling of the knots of Cauchy matrices, in sharp contrast to the impact of shift and scaling of the knots of Vandermonde matrices.

Theorem 49. *Assume a positive tolerance $\epsilon < 1$ and an $n \times n$ Cauchy matrix $C_{\mathbf{s},\mathbf{t}}$ with all $2n$ knots s_i and t_j lying on a circle on the complex plane. Then (cf. Remark 52) (i) $\alpha_\epsilon(C) = O(n\log(1/\epsilon))$, whereas (ii) $\beta_\epsilon(C) = O(n\log(n/\epsilon))$.*

Proof. Combine equations (25) and (26) to reduce the proof to the case where the $2n$ knots lie on the unit circle $\{z:\ |z|=1\}$. Then fix any complex a such that $|a|=1$ and recall that the function $\frac{z}{a} = 1 + \frac{2\sqrt{-1}}{z'-\sqrt{-1}}$ and its converse $z' = \frac{z+a}{z-a}\sqrt{-1}$ transform the real line into this unit circle and vice versa. Now write $s'_i = \frac{s_i+a}{s_i-a}\sqrt{-1}$ and $t'_j = \frac{t_j+a}{t_j-a}\sqrt{-1}$ for all i and j and obtain that all knots s'_i and t'_j are real, whereas $s_i = a(1+\frac{2\sqrt{-1}}{s'_i-\sqrt{-1}})$, $t_j = a(1+\frac{2\sqrt{-1}}{t'_j-\sqrt{-1}})$, $s_i - t_j = 2a\frac{s'_i-t'_j}{(s'_i-\sqrt{-1})(t'_j-\sqrt{-1})\sqrt{-1}}$, and consequently $\frac{1}{s_i-t_j} = \frac{u_iv_j}{s'_i-t'_j}$ for $u_i = \frac{\sqrt{-1}}{2a}(s'_i-\sqrt{-1})$ and $v_j = t'_j - \sqrt{-1}$. It follows that the Cauchy matrix $C = C_{\mathbf{s},\mathbf{t}}$ satisfies

$$C = \operatorname{diag}(\widehat{u}_i)_{i=0}^{n-1} C_{\mathbf{s}',\mathbf{t}'} \operatorname{diag}(v_j)_{j=0}^{n-1} \text{ for } \mathbf{s}' = (s'_i)_{i=0}^{n-1} \text{ and } \mathbf{t}' = (t'_j)_{j=0}^{n-1}. \tag{27}$$

Apply Theorem 9 to the matrix $C_{\mathbf{s}',\mathbf{t}'}$ and deduce that $\alpha_\epsilon(C) = O(n\log(1/\epsilon'))$ and $\beta_\epsilon(C) = O(n\log(1/\epsilon'))$ for $\epsilon' \le u\ v\ \epsilon$, $u = \max_{i=0}^{n-1}|u_i|$ and $v = \max_{j=0}^{n-1}|v_j|$. Recall that $|a|=1$ and deduce that $|u_i| \le 0.5(|s'_i|+1)$ and $|v_j| \le (|t'_j|+1)$ for all i and j. Recall that $|s_i| = |t_j| = 1$ and so $|s'_i| \le 2/|s_i-a|$ and $|t'_j| \le 2/|t_j-a|$ for all i and j. Choose a point a on the unit circle lying at the maximal distance $\delta \ge \delta_- = 2\sin(0.25\pi/n)$ from the set of the $2n$ knots $\{s_0,\dots,s_{n-1},t_0,\dots,t_{n-1}\}$. It follows that $\delta_- \ge 2/n$ for $n > 3$. Consequently $u \le 0.5(\frac{2}{\delta}+1) \le 0.5(n+1)$, $v \le \frac{2}{\delta}+1 \le n+1$, and $\epsilon' \le 0.5(\frac{2}{\delta}+1)^2\ \epsilon \le 0.5(n+1)^2\ \epsilon$ for $n>3$, which implies the bounds $\alpha_\epsilon(C)=O(n\log(n/\epsilon))$ and $\beta_\epsilon(C)=O(n\log(n/\epsilon))$.

To decrease the bound on $\alpha_\epsilon(C)$, partition the unit circle $\{z:\ |z|=1\}$ into three semi-open arcs $\mathcal{A}_h=\{\exp(\phi\sqrt{-1}):\ 2\pi h/3\le\phi<2\pi(h+1)/3\}$, each of length $2\pi/3$, for $h=0,1,2$, and write $C_h=(c_{i,j}^{(h)})_{i,j=0}^{n-1}$, $c_{i,j}^{(h)}=0$ if $s_i,t_j\in\mathcal{A}_h$, $c_{i,j}^{(h)}=\frac{1}{s_i-t_j}$ otherwise, for $h=0,1,2$. We can estimate $\alpha_\epsilon(C_h)$ by applying our previous argument, but now we choose $a=a(h)$ being the midpoint of the arc $\mathcal{A}_h$ and observe that in this case δ_- increases to 1, which implies that $\epsilon'\le 0.5(\frac{2}{\delta}+1)^2\ \epsilon\le 4.5\ \epsilon$ and $\alpha_\epsilon(C_h)=O(n\log(1/\epsilon))$ for $h=0,1,2$. Finally observe that $C=C_0+C_1+C_2$ and obtain that $\alpha_\epsilon(C)\le\sum_{h=0}^{2}\alpha_\epsilon(C_h)=O(n\log(1/\epsilon))$. □

We immediately extend the bound on $\alpha_\epsilon(C)$ to the case of Cauchy-like matrices M of part (c) of Theorem 14 with the knots on a line or a circle. Namely we combine equation (21) with part (i) of Theorem 49 and obtain that

$$\alpha_{\epsilon'}(M)=O(dn\log(1/\epsilon)) \tag{28}$$

where ϵ' and ϵ are linked by equation (21). To estimate $\beta_\epsilon(M)$ for such matrices M we first observe the bound of order $O(\log(1/\epsilon))$ on the ϵ-rank of the above matrices C_0, C_1, C_2 and $C_{\mathbf{s},\mathbf{t}}=C_0+C_1+C_2$. It follows that $\beta_\epsilon(C_{\mathbf{s},\mathbf{t}})=O(n\log(n)\log^2(1/\epsilon))$ (cf. Theorem 43). Apply the techniques of Section 9.4 to extend this estimate to Cauchy-like matrices M as follows.

Corollary 50. *Suppose M is a Cauchy-like matrix of part (c) of Theorem 14 having its knots on a line or a circle. Then $\alpha_\epsilon(M)=O(nd\log(1/\epsilon'))$ and $\beta_{\epsilon''}(M)=O(nd^2\log(n)\log^2(1/\epsilon))$ for ϵ' and ϵ'' defined in Section 9.4.*

Finally recall the canonical DFT-based transforms of the matrices of the classes $\mathcal{T}$ and $\mathcal{H}$ into Cauchy-like matrices (see Theorem 26) and extend the corollary to obtain the following result.

Corollary 51. *Suppose M is a Toeplitz-like or Hankel-like matrix of parts (t) or (h) of Theorem 14. Then $\beta_{\epsilon''}(M)=O(nd^2\log(n)\log^2(1/\epsilon))$ for ϵ'' and ϵ' defined in Section 9.4 and for $d\le 2$ in the case of Toeplitz and Hankel matrices M.*

Remark 52. The transform of Cauchy matrices based on the function $\frac{az+b}{z-h}=a+\frac{b+ah}{z-h}$ is numerically unstable for many values of the parameters a, b and h, but we avoid numerical problems by employing the numerically stable algorithms of [CGS07], [X12], [XXG12], and [XXCB] and applying the functional transform just to bound the numerical ranks of the admissible blocks of Cauchy matrices involved into the computations. Then application of Theorem 43 for $\rho=O(\log(1/\epsilon))$ still yields the cost bounds $\alpha_{\epsilon'}(M)=O(nd\log(1/\epsilon))$ and $\beta_{\epsilon''}(M)=O(nd^2\log(n)\log^2(1/\epsilon))$.

9.9 ϵ-ranks of admissible blocks and the impact on implementation

Our proof of Theorem 43 is constructive, that is we can readily compute the centers c_q and the admissible blocks $\widehat{N}_q$ of bounded ranks throughout the merging process, and then we can apply the algorithms of the previous section. In practice one should avoid a large part of these computations, however, by following the papers [CGS07], [X12], [XXG12], and [XXCB]. They bypass the computation of the centers c_q and immediately compute the HSS generators for the admissible blocks $\widehat{N}_q$, defined by HSS trees. The length of the generators can be chosen equal to the available upper bound ρ on the numerical ranks of these blocks or can be adapted empirically. Theorem 43 implies that the computational cost bounds $\alpha_\epsilon(M)$ and $\beta_\epsilon(M)$ are proportional to ρ and ρ^2, respectively, and thus decrease as the numerical rank ρ decreases. If bound (20) on the ϵ-rank decreases to $\rho=O(\log(1/\epsilon))$ (cf. our Remark 40), then the complexity bounds of Theorem 43 decrease to the level $\alpha_\epsilon(C)=O(n\log(1/\epsilon)\log(n))$ and $\beta_\epsilon(C)=O(n\log^2(1/\epsilon)\log(n))$. By virtue of Corollary 51 this is the case for the inputs from the classes $\mathcal{T}$ and $\mathcal{H}$, including the case of Toeplitz and Hankel inputs, where such bounds have been empirically observed in [XXG12].

10 Conclusions

The techniques of the transformation of matrix structures based on displacement representation go back to [P90], with surprising algorithmic applications explored since 1995. At first we revisited these techniques covering them comprehensively. We simplified their study by employing Sylvester's (rather than Stein's) displacements and the techniques for operating with them from [P00] and [P01, Section 1.5]. Then we covered some fast numerically stable approximation algorithms based on combining these transformations with another link among distinct classes of structured matrices, namely among Cauchy and HSS matrices, the latter matrices appeared inthe study of the FMM (that is the Fast Multipole Method). These efficient algorithms approximate the solution of a nonsingular Toeplitz or Toeplitz-like linear system of equations in nearly linear (versus classical cubic) arithmetic time [MRT05], [CGS07], [XXG12], and [XXCB]. Our analysis of these algorithms revealed their additional power, and we extended them to support nearly linear arithmetic time bounds (versus known quadratic) for the approximation of the matrix-by-vector products of Vandermonde, transposed Vandermonde, and CV matrices, the latter ones being a subclass of Cauchy matrices and for approximate solution of nonsingular linear systems of equations with these matrices. We noted some potential numerical limitations for the application of our transformations to the latter task of solving linear systems of equations and for our algorithmic transformations of Section 9.5 from CV matrices to Cauchy matrices with any set of knots. We observed no such limitations, however, in extension of our results to the matrices of the classes $\mathcal{V}$, $\mathcal{V}^T$, and $\mathcal{CV}$, and we further accelerated a little the cited numerical approximation algorithms for Toeplitz linear systems by combining the algorithms of [DGR96] for polynomial evaluation with functional transformations of matrix structures.

At this point natural research challenges include (i) the search for new links and new transformations among various classes of structured matrices towards significant algorithmic applications (possibly by combining the displacement and functional transformations of matrix structures with the approximation techniques of the FMM) and (ii) the refinement of the presented algorithms. Our arithmetic time bounds for (l, u)-HSS computations exceed the bounds for similar computations with banded matrices by logarithmic factors, and one may try to close these gaps by applying the advanced techniques of the FMM. A more specific idea towards a specific goal is the combination of equation (12) and Example 46 in order to accelerate approximate solution of Vandermonde linear systems to the level achieved for multiplication of a transposed Vandermonde matrix by a vector.

Even the acceleration to the level of Theorem 9, however, would support substantially inferior Boolean complexity estimates (with the excess by a factor of $\log(1/\epsilon)$) compared to the algorithms of [BP87, Main Theorem], [K98, Theorem 3.9 and Section 5.3], and [PT13] for high precision polynomial evaluation and interpolation and consequently for the related Vandermonde and Cauchy matrix computations. Equation (5) may help to extend the latter algorithms to high precision multiplication of a Cauchy matrix by a vector and solving a Cauchy linear system of equations. Such progress may eventually become of interest for numerical computations as well, in the case of sufficient support from the field of Computer Arithmetic (cf. [P91]).

Acknowledgements: Our research has been supported by NSF Grant CCF–1116736 and PSC CUNY Awards 64512–0042 and 65792–0043. We also wish to thank the reviewers for thoughtful valuable comments.

References

[B10] S. Börm, *Efficient Numerical Methods for Non-local Operators: $\mathcal{H}^2$-Matrix Compression, Algorithms and Analysis*, European Math. Society, 2010.

[BA80] R. R. Bitmead, B. D. O. Anderson, Asymptotically Fast Solution of Toeplitz and Related Systems of Linear Equations, *Linear Algebra and Its Applications*, **34**, 103–116, 1980.

[BEGO08] T. Bella, Y. Eidelman, I. Gohberg, V. Olshevsky, Computations with Quasiseparable Polynomials and Matrices, *Theoretical Computer Science*, *Special Issue on Symbolic–*

Numerical Algorithms (D. A. Bini, V. Y. Pan, and J. Verschelde editors), **409**, **2**, 158–179, 2008.

[BF00] D. A. Bini, G. Fiorentino, Design, Analysis, and Implementation of a Multiprecision Polynomial Rootfinder, *Numer. Algs.*, **23**, 127–173, 2000.

[BM01] D. A. Bini, B. Meini, Approximate Displacement Rank and Applications, in *AMS Conference "Structured Matrices in Operator Theory, Control, Signal and Image Processing"*, Boulder, 1999 (edited by V. Olshevsky), *Amer. Math. Soc.*, 215–232, Providence, RI, 2001.

[BP87] D. Bini, V. Y. Pan, A Logarithmic Boolean Time Algorithm for Parallel Polynomial Division, *Information Processing Letters*, **24**, 233–237 (1987).

[BP94] D. Bini, V. Y. Pan, *Polynomial and Matrix Computations, Volume 1: Fundamental Algorithms*, Birkhäuser, Boston, 1994.

[BY13] L. A. Barba, R. Yokota, How Will the Fast Multipole Method Fare in Exascale Era? *SIAM News*, **46**, **6**, 1–3, July/August 2013.

[CDGLP06] S. Chandrasekaran , P. Dewilde , M. Gu , W. Lyons , T. Pals, A Fast Solver for HSS Representations via Sparse Matrices, *SIAM J. Matrix Anal. Appl.*, **29, 1**, 67–81, 2006.

[CGR98] J. Carrier, L. Greengard, V. Rokhlin, A Fast Adaptive Algorithm for Particle Simulation, *SIAM J. Scientific Computing*, **9**, 669–686, 1998.

[CGS07] S. Chandrasekaran, M. Gu, X. Sun, J. Xia, J. Zhu, A Superfast Algorithm for Toeplitz Systems of Linear Equations, *SIAM J. Matrix Anal. Appl.*, **29**, 1247–1266, 2007.

[CPW74] R. E. Cline, R. J. Plemmons, and G. Worm, Generalized Inverses of Certain Toeplitz Matrices, *Linear Algebra and Its Applications,* **8**, 25–33, 1974.

[DGR96] A. Dutt, M. Gu, V. Rokhlin, Fast Algorithms for Polynomial Interpolation, Integration, and Differentiation, *SIAM Journal on Numerical Analysis*, **33, 5**, 1689–1711, 1996.

[DV98] P. Dewilde and A. van der Veen, *Time-Varying Systems and Computations*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[EG02] Y. Eidelman, I. Gohberg, A Modification of the Dewilde–van der Veen Method for Inversion of Finite Structured Matrices, *Linear Algebra and Its Applications*, **343**, 419–450, 2002.

[EGH13a] Y. Eidelman, I. Gohberg, I. Haimovici, *Separable Type Representations of Matrices and Fast Algorithms. Volume 1: Basics. Completion problems. Multiplication and Inversion Algorithms*, Birkhauser, 2013.

[EGH13b] Y. Eidelman, I. Gohberg, I. Haimovici, *Separable Type Representations of Matrices and Fast Algorithms. Volume 2: Eigenvalue Methods*, Birkhauser, 2013.

[G98] M. Gu, Stable and Efficient Algorithms for Structured Systems of Linear Equations, *SIAM J. Matrix Anal. Appl.*, **19**, 279–306, 1998.

[GE96] M. Gu, S. C. Eisenstat, Efficient Elgorithms for Computing a Strong Rank-Revealing QR Factorization, *SIAM Journal on Scientific Computing*, **17**, 848–869, 1996.

[GI88] W. Gautschi, G. Inglese, Lower Bounds for the Condition Number of Vandermonde Matrices, *Numerische Mathematik*, **52**, 241–250, 1988.

[GKK85] I. Gohberg, T. Kailath, I. Kailath, Linear Complexity Algorithms for Semiseparable Matrices, *Integral Equations and Operator Theory*, **8, 6**, 780–804, 1985.

[GKO95] I. Gohberg, T. Kailath, V. Olshevsky, Fast Gaussian Elimination with Partial Pivoting for Matrices with Displacement Structure, *Mathematics of Computation*, **64**, 1557–1576, 1995.

[GO94] I. Gohberg, V. Olshevsky, Complexity of Multiplication with Vectors for Structured Matrices, *Linear Algebra and Its Applications*, **202**, 163–192, 1994.

[GL96] G. H. Golub, C. F. Van Loan, *Matrix Computations* (third edition), The Johns Hopkins University Press, Baltimore, Maryland, 1996.

[GR87] L. Greengard, V. Rokhlin, A Fast Algorithm for Particle Simulation, *Journal of Computational Physics*, **73**, 325–348, 1987.

[H95] G. Heinig, Inversion of Generalized Cauchy Matrices and the Other Classes of Structured Matrices, *Linear Algebra for Signal Processing, IMA Volume in Mathematics and Its Applications*, **69**, 95–114, Springer, 1995.

[K98] P. Kirrinnis, Polynomial Factorization and Partial Fraction Decomposition by Simultaneous Newton's Iteration, *J. of Complexity*, **14**, 378–444, 1998.

[KKM79] T. Kailath, S. Y. Kung, M. Morf, Displacement Ranks of Matrices and Linear Equations, *Journal of Mathematical Analysis and Applications*, **68, 2**, 395–407, 1979.

[KZ08] S. Köhler, M. Ziegler, On the Stability of Fast Polynomial Arithmetic, *Proc. 8-th Conference on Real Numbers and Computers* (J.D. Bruguera and M. Daumas, eds.), 147–156, 2008.

[LRT79] R. J. Lipton, D. Rose, R. E. Tarjan, Generalized Nested Dissection, *SIAM J. on Numerical Analysis*, **16**, **2**, 346–358, 1979.

[M80] M. Morf, Doubling Algorithms for Toeplitz and Related Equations, *Proceedings of IEEE International Conference on ASSP*, 954–959, IEEE Press, Piscataway, New Jersey, 1980.

[MRT05] P. G. Martinsson, V. Rokhlin, M. Tygert, A Fast Algorithm for the Inversion of General Toeplitz Matrices, *Comput. Math. Appl.*, **50**, 741–752, 2005.

[P90] V. Y. Pan, On Computations with Dense Structured Matrices, *Math. of Computation*, **55, 191**, 179–190, 1990. Proceedings version in *Proc. Intern. Symposium on Symbolic and Algebraic Computation (ISSAC'89)*, 34–42, ACM Press, NY, 1989.

[P91] D. Priest, Algorithms for arbitrary precision floating point arithmetic, In P. Kornerup and D. Matula, eds., *Proc. 10th Symp. Computer Arithmetic*, 132–145, IEEE Computer Society Press, Los Angeles, California, 1991.

[P93] V. Y. Pan, Parallel Solution of Sparse Linear and Path Systems, in *Synthesis of Parallel Algorithms* (J.H. Reif editor), Chapter 14, pp. 621-678, Morgan Kaufmann publishers, San Mateo, California, 1993.

[P95] V. Y. Pan, An Algebraic Approach to Approximate Evaluation of a Polynomial on a Set of Real Points, *Advances in Computational Mathematics*, **3**, 41–58, 1995.

[P00] V. Y. Pan, Nearly Optimal Computations with Structured Matrices, *Proceedings of 11th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'2000)*, 953–962, ACM Press, New York, and SIAM Publications, Philadephia, 2000.

[P01] V. Y. Pan, *Structured Matrices and Polynomials: Unified Superfast Algorithms*, Birkhäuser/Springer, Boston/New York, 2001.

[P11] V. Y. Pan, Nearly Optimal Solution of Rational Linear Systems of Equations with Symbolic Lifting and Numerical Initialization, *Computers and Mathematics with Applications*, **62**, 1685–1706, 2011.

[Pa] V. Y. Pan, Fast Approximation Algorithms for Cauchy Matrices, Polynomials and Rational Functions, to appear.

[PR93] V. Y. Pan, J. Reif, Fast and Efficient Parallel Solution of Sparse Linear Systems, *SIAM J. on Computing*, **22**, **6**, 1227–1250, 1993.

[PRT92] V. Y. Pan, J. H. Reif, S. R. Tate, The Power of Combining the Techniques of Algebraic and Numerical Computing: Improved Approximate Multipoint Polynomial Evaluation and Improved Multipole Algorithms, *33th Annual IEEE Symposium on Foundations of Computer Science (FOCS'92)*, 703–713, IEEE Computer Societty Press, 1992.

[PSLT93] V. Y. Pan, A. Sadikou, E. Landowne, O. Tiga, A New Approach to Fast Polynomial Interpolation and Multipoint Evaluation, *Computers and Math. (with Applications)*, **25, 9**, 25–30, 1993.

[PT13] V. Y. Pan, E. P. Tsigaridas, On the Boolean Complexity of the Real Root Refinement, preprint 2013. Proc. version in in *Proc. Intern. Symposium on Symbolic and Algebraic Computation (ISSAC 2013)*, (M. Kauers ed.), 299–306, Boston, MA, June 2013, ACM Press, New York, 2013.

[PW03] V. Y. Pan, X. Wang, Inversion of Displacement Operators, *SIAM J. on Matrix Analysis and Applications*, **24, 3**, 660–677, 2003.

[PZHY97] V. Y. Pan, A. Zheng, X. Huang, Y. Yu, Fast Multipoint Polynomial Evaluation and Interpolation via Computation with Structured Matrices, *Annals of Numerical Math.*, **4**, 483–510, 1997.

[R85] V. Rokhlin, Rapid Solution of Integral Equations of Classical Potential Theory, *Journal of Computational Physics*, **60**, 187–207, 1985.

[R06] G. Rodriguez, Fast Solution of Toeplitz- and Cauchy-like Least Squares Problems, *SIAM J. Matrix Analysis and Applications*, **28, 3**, 724–748, 2006.

[S98] G. W. Stewart, *Matrix Algorithms, Vol I: Basic Decompositions*, SIAM, 1998.

[T00] E.E. Tyrtyshnikov, Incomplete Cross-Approximation in the Mosaic-Skeleton Method, *Computing*, **64**, 367–380, 2000.

[VVGM05] R. Vandebril, M. Van Barel, G. Golub, N. Mastronardi, A Bibliography on Semiseparable Matrices, *Calcolo*, **42, 3–4,** 249–270, 2005.

[VVM07] R. Vandebril, M. Van Barel, N. Mastronardi, *Matrix Computations and Semiseparable Matrices: Linear Systems* (Volume 1), The Johns Hopkins University Press, Baltimore, Maryland, 2007.

[VVM08] R. Vandebril, M. Van Barel, N. Mastronardi, *Matrix Computations and Semiseparable Matrices: Eigenvalue and Singular Value Methods* (Volume 2), The Johns Hopkins University Press, Baltimore, Maryland, 2008.

[X12] J. Xia, On the Complexity of Some Hierarchical Structured Matrix Algorithms, *SIAM J. Matrix Anal. Appl.*, **33**, 388–410, 2012.

[X13] J. Xia, Randomized sparse direct solvers, *SIAM J. Matrix Anal. Appl.*, **34**, 197–227, 2013.

[XXCB] Y. Xi, J. Xia, S. Cauley, V. Balakrishnan, Superfast and Stable Structured Solvers for Toeplitz Least Squares via Randomized Sampling, submitted, *SIAM J. Matrix Anal. Appl.*

[XXG12] J. Xia, Y. Xi, M. Gu, A Superfast Structured Solver for Toeplitz Linear Systems via Randomized Sampling, *SIAM J. Matrix Anal. Appl.*, **33**, 837–858, 2012.