2013

TR-2013004: Transformations of Matrix Structures Work Again

Victor Y. Pan

Transformations of Matrix Structures Work Again

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

In [P90] we proposed to employ Vandermonde and Hankel multipliers to transform into each other the matrix structures of Toeplitz, Hankel, Vandermonde and Cauchy types as a means of extending any successful algorithm for the inversion of matrices having one of these structures to inverting the matrices with the structures of the three other types. Surprising power of this approach has been demonstrated in a number of works, which culminated in ingeneous numerically stable algorithms that approximated the solution of a nonsingular Toeplitz linear system in nearly linear (versus previuosly cubic) arithmetic time. We first revisit this powerful method, covering it comprehensively, and then specialize it to yield a similar acceleration of the known algorithms for computations with matrices having structures of Vandermonde or Cauchy types. In particular we arrive at numerically stable approximate multipoint polynomial evaluation and interpolation in nearly linear time, by using $O(bn\log^h n)$ flops where $h=1$ for evaluation, $h=2$ for interpolation, and 2^{-b} is the relative norm of the approximation errors.

Keywords: Transforms of matrix structures, Vandermonde matrices, Cauchy matrices, Multipole method, HSS matrices, Polynomials, Rational functions, Multipoint evaluation, Interpolation

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

1 Introduction

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 a number of their further attractive properties:

- Compact compressed representation through a small number of parameters, typically $O(n)$ parameters in the case of $n\times n$ matrices
- Simple expressions for the inverse through the solutions of a small number of linear systems of equations wherever the matrix is invertible
- 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

Table 1: Four classes of structured matrices

.

Toeplitz matrices $T = (t_{i-j})_{i,j=1}^{n}$	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 $C = C_{\mathbf{s},\mathbf{t}} = \left(\frac{1}{s_i - t_j}\right)_{i,j=1}^{n}$
$\begin{pmatrix} 1 & s_1 & \cdots & s_1^{n-1} \\ 1 & s_2 & \cdots & s_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & s_n & \cdots & s_n^{n-1} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{s_1-t_1} & \cdots & \frac{1}{s_1-t_n} \\ \frac{1}{s_2-t_1} & \cdots & \frac{1}{s_2-t_n} \\ \vdots & & \vdots \\ \frac{1}{s_n-t_1} & \cdots & \frac{1}{s_n-t_n} \end{pmatrix}$

Extensive and 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 structures into each other. For example, $\mathcal{T}\mathcal{H} = \mathcal{H}\mathcal{T} = \mathcal{H}$, $\mathcal{H}\mathcal{H} = \mathcal{T}$, and V^TV is a Hankel matrix. The paper [P90] showed that *this technique enables 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.* We cover this technique comprehensively and simplify its presentation versus [P90] because instead of the the Stein displacements $M - AMB$ in [P90] we employ the Sylvester displacements $AM - MB$ and the machinery of operating with them from [P00] and [P01, Section 1.5].

The proposed structure transforms are simple but have surprising power where the transform links matrix classes having distinct features. For example, the matrix structure of Cauchy type is invariant in row and column interchange (in contrast to the structures of Toeplitz and Hankel types) and enables expansion of the matrix entries into Loran's series (unlike the structures of the three other types). Exploiting these distinctions has lead to dramatic acceleration of the known numerically stable algorithms for Toeplitz and Toeplitz-like linear systems of equations by means of their transformation into Cauchy-like matrices and exploiting the above properties of these matrices.

Their invariance to row interchange enabled numerically stable solution in quadratic rather than cubic time in [GKO95], [G98], [R06], but the paper [MRT05] (cf. also [CGS07], [XXG12], and [XXCB]) has instead exploited the Loran's expansion of the entries of the basic Cauchy matrices to obtain their close approximation by HSS matrices. ("HSS" is the acronym for "hierarchically semiseparable".) This structure of a distinct type enabled application of the *Multipole/HSS* powerful techniques, and the resulting numerically stable algorithms approximate the solution of a nonsingular Toeplitz linear system of equations in nearly linear (and thus nearly optimal) arithmetic time. The intensive work in [XXG12] and [XXCB] on extension, refinement and implementation of the algorithms has already made them quite attractive for the users.

Similar advance has not been achieved, however, for the computations with matrices having structures of Vandermonde or Cauchy types. All the cited papers on Toeplitz computations share their basic displacement map, which is a specialization of our general class of the transformations of matrix structures derived from [P90] (see our comments at the end of Section 5.3). The map transforms the matrices with the structure of Toeplitz type into the matrices of the subclasses of the class $\mathcal{C}$ linked to FFT and defined by the knot sets $\{s_1, \dots, s_n, t_1, \dots, t_n\}$ equally spaced on the unit circle $\{z : \ |z| = 1\}$ of the complex plane. This covers the structures of Toeplitz but not Vandermonde and Cauchy types.

In our present paper we specify the subclass of *CV* and *CV-like* matrices, which are the Cauchy and Cauchy-like matrices, respectively, having at least one (but not necessarily both) of their two basic knot sets $\{s_1, \dots, s_n\}$ or $\{t_1, \dots, t_n\}$ equally spaced on the unit circle $\{z : \ |z| = 1\}$. These are precisely the Cauchy and Cauchy-like matrices that have FFT-type structured transforms into the matrices of the class $\mathcal{V}$ or their transposes. Under this framework our main technical step is an extension of the algorithms of [MRT05], [CGS07], and [XXG12] to proving that all CV and CV-like matrices can be closely appoximated by HSS matrices. As soon as such an approximation is available, one just needs to apply the Multipole method to the HSS matrices to obtain numerically stable approximation algorithms that run in nearly linear time for our tasks for CV matrices, versus quadratic time of the known algorithms. By applying the FFT-based structured transforms between matrices with the structures of CV and Vandermonde types, we readily extend these results to the matrices of the latter class and consequently to the problems of multipoint evaluation and interpolation for polynomials.

The new algorithms approximate within relative error norm bound 2^{-b} the product of an $n \times n$ CV matrix by a vector by using $O(bn \log n)$ flops and the solution of a nonsingular CV linear system of n equations by using $O(bn \log^2 n)$ flops. FFT-based structured transforms extend these algorithms and complexity bounds to computations with Vandermonde matrices and to approximate multipoint evaluation and interpolation for polynomials. The resulting nearly linear time bounds are nearly optimal, but still seem to be overly pessimistic, in view of the results of the extensive tests in [XXG12] for the similar HSS computations (see our Remark 36). The cited results are readily extended to CV-like matrices and consequently to the matrices having structure of Vandermonde type.

Various extensions, ameliorations, refinements, and nontrivial specializations of the proposed methods can be interesting. Most valuable would be new transforms among various new classes of structured matrices, with significant algorithmic applications. At the end of Section 9 we sketch a natural extension of our techniques to the general class of Cauchy and Cauchy-like matrices, but indicate that this generally complicates the control over the output errors. It can be interesting that even a very crude variant of our techniques (which proceeds with a limited use of the HSS algorithms) still accelerates the known numerical algorithms for multipoint polynomial evaluation by a factor of $\sqrt{n/\log n}$ (see Remark 34).

For a sample further application, recall that the current best package of subroutines for polynomial root-finding, MPSolve, is reduced essentially to recursive application of the Ehrlich–Aberth algorithm, and consequently to recursive numerical multipoint polynomial evaluation. For this task MPSolve uses a quadratic time algorithm, which is the users' current choice. So our present acceleration from quadratic to a nearly linear time can be translated into the same acceleration of MPSolve.

We organize our pesentation 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, recall the class of 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 approximations of these matrices in Section 8 and to approximate the products of these matrices and their inverses by a vector in Section 9. 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.

2 Some definitions and basic facts

Hereafter "flop" stands for "arithmetic operation"; 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 (complex conjugate) transpose, respectively. We write M^{-T} for $(M^T)^{-1} = (M^{-1})^T$.

$(B_1 \mid \dots \mid B_n)$ denotes a $1 \times n$ block matrix with the blocks $B_1, \dots, B_n$. $\mathrm{diag}(B_1, \dots, B_n) = \mathrm{diag}(B_j)_{j=1}^n$ is an $n \times n$ block diagonal matrix with the diagonal blocks $B_1, \dots, B_n$. In the case of scalar blocks $s_1, \dots, s_n$ we arrive at a vector $\mathbf{s} = (s_j)_{j=1}^n$ and an $n \times n$ diagonal matrix $D_{\mathbf{s}} = \mathrm{diag}(\mathbf{s}) = \mathrm{diag}(s_j)_{j=1}^n$ with the diagonal entries $s_1, \dots, s_n$.

The n coordinate vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$ of a dimension n form the $n \times n$ identity matrix $I_n = (\mathbf{e}_1 \mid \dots \mid \mathbf{e}_n)$ and the $n \times n$ reflection matrix $J_n = (\mathbf{e}_n \mid \dots \mid \mathbf{e}_1)$. $J_n = J_n^T = J_n^{-1}$. We write I and J where the matrix size is not important or is defined by context.

Preprocessors. For three nonsingular matrices P, M, and N and a vector $\mathbf{b}$, the equations

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

reduce the inversion of the matrix M and the solution of a linear system of equations $M\mathbf{x} = \mathbf{b}$ to the inversion of the product PMN and the solution of the linear system $PMN\mathbf{y} = P\mathbf{b}$, respectively. For some important classes of matrices M this preprocessing can simplify dramatically the inversion of a matrix and the solution of a linear system of equations.

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, which is the shortest for $l = r$.

Theorem 1. *(Cf. [BA80], [M80], [P93], [GE96], [P01, Section 4.6.2].) Given a generator of a length l for an $n \times n$ matrix M having a rank r, $r \leq l \leq n$, it is sufficient to use $O(l^2 n)$ flops to compute a generator of length r for the matrix M.*

Norm, conditioning, orthogonality, numerical rank, a perturbation norm bound. $||M|| = ||M^H|| = ||M||_2$ is the (Euclidean) 2-norm of a matrix M.

For a fixed tolerance τ the minimum rank of matrices in the τ-neighborhood of a matrix M is said to be its *τ-rank*. The *numerical rank* of a matrix is its τ-rank for a small positive τ. A matrix is called ill conditioned if it has a close neighbor of a smaller rank or eqiuivalently if its rank exceeds its numerical rank. Otherwise it is called *well conditioned*. If a matrix M is ill conditioned, one must compute its inverse and the solution of a linear system $M\mathbf{x} = \mathbf{f}$ with a high precision to ensure meaningful output for these problems, but *not for multiplication* by a vector.

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 constant c. Such a matrix U has full rank and is very well conditioned: its distance to the closest matrix of a smaller rank is equal to $||U|| = 1$.

Theorem 2. *(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 $|||(M+E)^{-1} - M^{-1}|| \leq 0.5||M^{-1}||$ if $\theta \leq 1/3$.*

2.2 The classes of Toeplitz, Hankel, Vandermonde and Cauchy matrices, some subclasses and factorizations, polynomial evaluation and interpolation

For larger 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, as is the case with the Vandermonde matrices Ω and Ω^H below, which are unitary up to scaling by $\frac{1}{\sqrt{n}}$.

DFT and DFT-based matrices. (See [BP94, Sections 1.2, 3.4].) Write $\omega_n = \exp(\frac{2\pi}{n}\sqrt{-1})$ to denote a primitive nth root of 1. Its powers $1, \omega_n, \dots, \omega_n^{n-1}$ are equally spaced on the unit circle $\{z : |z| = 1\}$. Let $\Omega = \Omega_n = (\omega_n^{ij})_{i,j=0}^{n-1}$ denote the $n \times n$ matrix of *DFT*, that is of the *discrete Fourier transform* at n points. Ω and Ω^H are quasiunitary, whereas $\frac{1}{\sqrt{n}}\Omega$ and $\frac{1}{\sqrt{n}}\Omega^H$ and $\Omega^{-1} = \frac{1}{n}\Omega^H$ are unitary matrices, because $\Omega\Omega^H = nI$. See, e.g., [BP94, Sections 1.2 and 3.4] on a proof of the following theorem and on the numerical stability of the supporting algorithms.

Theorem 3. *For any vector* $\mathbf{v} = (v_i)_{i=1}^n$ *one can compute the vectors* $\Omega\mathbf{v}$ *and* $\Omega^{-1}\mathbf{v}$ *by using* $O(n \log n)$ *flops. If* $n = 2^k$ *is a power of 2, then one can compute the vectors* $\Omega\mathbf{v}$ *and* $\Omega^{-1}\mathbf{v}$ *by using* $0.5n\log_2 n$ *and* $0.5n\log_2 n + n$ *flops, respectively.*

Cauchy and Vandermonde matrices and polynomial evaluation and interpolation. (See Table 1 and [P01, Chapters 2 and 3].) It holds that

$$C_{\mathbf{s},\mathbf{t}} = -C_{\mathbf{t},\mathbf{s}}^T, \tag{2}$$

$$C_{\mathbf{s},\mathbf{t}} = \operatorname{diag}(t(s_i)^{-1})_{i=1}^n V_{\mathbf{s}} V_{\mathbf{t}}^{-1} \operatorname{diag}(t'(t_i))_{i=1}^n \tag{3}$$

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

Equation (3) expresses a Cauchy matrix $C_{\mathbf{s},\mathbf{t}}$ through the Vandermonde matrix $V_{\mathbf{s}}$, the inverse $V_{\mathbf{t}}^{-1}$ of the Vandermonde matrix $V_{\mathbf{t}}$, the coefficients of the auxiliary polynomial $t(x)$ defined by its roots $t_0, \dots, t_{n-1}$, and the values of this polynomial and its derivative $t'(x)$, each at n points. Part (i) of the following simple theorem states that *polynomial multipoint evaluation and interpolation* with the knots $s_1, \dots, s_n$ are equivalent to multiplication of the Vandermonde matrix $V_{\mathbf{s}}$ by the coefficient vector of the polynomial and the solution of the associated linear system of equations with this matrix, respectively. Part (ii) of the theorem shows shows equivalence of rational multipoint evaluation and interpolation to the similar equations for Cauchy (rather than Vandermonde) matrix. Part (iii) of the theorem shows that the reconstruction of the polynomial coefficients from the roots can be reduced to polynomial interpolation and consequently to solving a Vandermonde linear system of equations. on these links and similar links of rational multipoint evaluation and interpolation to Cauchy matrices. Together with equation (3), the theorem also links multipoint evaluation and interpolation for polynomials to the same tasks for rational functions (cf. [P01, Chapter 3]).

Theorem 4. *(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* $v(x) = \sum_{j=1}^n \frac{u_j}{x - t_j}$ *with* n *distinct poles* $t_1, \dots, t_n$ *and for* n *distinct scalars* $s_1, \dots, s_n$*, write* $\mathbf{s} = (s_i)_{i=1}^n$, $\mathbf{t} = (t_j)_{j=1}^n$, $\mathbf{u} = (u_j)_{j=1}^n$, $\mathbf{v} = (v_i)_{i=1}^n$*. Then the equations* $v_i = v(s_i)$, $i = 1, \dots, n$ *hold if and only if* $C_{\mathbf{s},\mathbf{t}}\mathbf{u} = \mathbf{v}$*. (iii) The equation* $\prod_{i=0}^{n-1}(x - t_i) = x^n + v(x)$*, for* $v(x) = \sum_{i=0}^{n-1} t_i x^i$ *and for* n *distinct knots* $t_0, \dots, t_{n-1}$*, is equivalent to the linear system of* n *equations,* $v(t_i) = -t_i^n$ *for* $i = 0, \dots, n-1$.

Theorem 5. *(i)* $\det(V) = \prod_{i<k}(s_i - s_k)$ *and* $\det(C) = \prod_{i<j}(s_i - s_j)(t_i - t_j)/\prod_{i,j}(s_i - t_j)$*, and so the matrices* V *and* C *of Table 1 are nonsingular where all scalars* $s_1, \dots, s_n, t_1, \dots, t_n$ *are distinct. (ii) A row interchange preserves both Vandermonde and Cauchy structures. A column interchange preserves Cauchy structure.*

Next we will specify a subclass of Cauchy matrices most closely linked to Vandermonde and transposed Vandermonde matrices (cf. Definition 6). At first write

$$V_f = ((f\omega_n^{i-1})^{j-1})_{i,j=1}^n = \Omega \operatorname{diag}(f^{j-1})_{j=1}^n, \tag{4}$$

$$C_{\mathbf{s},f} = \left(\frac{1}{s_i - f\omega_n^{j-1}}\right)_{i,j=1}^n, \; C_{e,\mathbf{t}} = \left(\frac{1}{e\omega_n^{i-1} - t_j}\right)_{i,j=1}^n, \; C_{e,f} = \left(\frac{1}{e\omega_n^{i-1} - f\omega_n^{j-1}}\right)_{i,j=1}^n \tag{5}$$

for two distinct scalars e and $f \neq 0$. Then observe that $\Omega = V_1 = V_{\mathbf{s}}$ for $\mathbf{s} = (\omega_n^{i-1})_{i=1}^n$, $\Omega^H = V_{\mathbf{t}}$ for $\mathbf{t} = (\omega_n^{1-i})_{i=1}^n$ (so both Ω and $\Omega^H$are Vandermonde matrices), $C_{\mathbf{s},f} = C_{\mathbf{s},\mathbf{t}}$, $C_{e,\mathbf{t}} = C_{\mathbf{s},\mathbf{t}}$, $C_{e,f} = C_{\mathbf{s},\mathbf{t}}$ where $\mathbf{s} = (e\omega_n^{j-1})_{i,j=1}^n$ and/or $\mathbf{t} = (f\omega_n^{j-1})_{i,j=1}^n$, and *the matrices* V_f *are quasiunitary* where $|f| = 1$.

Now let $\mathbf{t} = (f\omega_n^{j-1})_{j=1}^n$, obtain $t(x) = x^n - f^n$, $t(x) = nx^{n-1}$, $t(s_i) = s_i^n - f^n$, $t'(t_i) = nf^{n-1}\omega_n^{1-i}$ for all i, and $nV_f^{-1} = \operatorname{diag}(f^{1-i})_{i=1}^n \Omega^H$, substitute into (3), and obtain

$$C_{\mathbf{s},f} = \operatorname{diag}\left(\frac{f^{n-1}}{s_i^n - f^n}\right)_{i=1}^n V_{\mathbf{s}} \operatorname{diag}(f^{1-i})_{i=1}^n \Omega^H \operatorname{diag}(\omega_n^{1-i})_{i=1}^n. \tag{6}$$

If in addition $\mathbf{s} = (e\omega_n^{i-1})_{i=1}^n$, then $s_i^n = e^n$ for all i and $V_{\mathbf{s}} = V_f$. Substitute into (6) and obtain

$$C_{e,f} = n\frac{f^{n-1}}{e^n - f^n}\Omega \operatorname{diag}((e/f)^{i-1})_{i=1}^n \Omega^H \operatorname{diag}(\omega^{1-i})_{i=1}^n. \tag{7}$$

Definition 6. Hereafter we refer to the matrices V_f, $C_{\mathbf{s},f}$, $C_{e,\mathbf{t}}$, and $C_{e,f}$ for all 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.

Equations (4) and (7) link the DFT-based matrices to the DFT matrix Ω. Similarly to this matrix they have their basic sets of knots $\mathbb{S} = \{s_1, \ldots, s_n\}$ and $\mathbb{T} = \{t_1, \ldots, t_n\}$ equally spaced on the unit circle $\{z : \ |z| = 1\}$. Equations (6) link the CV matrices to Vandermonde matrices $V_{\mathbf{s}}$ and $V_{\mathbf{t}}$, respectively. Combine equation (4) and (7) with Theorem 3 to obtain the following results.

Theorem 7. *$O(n \log n)$ flops are sufficient to compute the product $M\mathbf{f}$ of a DFT-based Vandermonde or Cauchy $n \times n$ matrix M and a vector $\mathbf{f}$. If the matrix M is nonsingular, then $O(n \log n)$ flops are also sufficient to compute the solution $\mathbf{x}$ to a linear system of n equations $M\mathbf{x} = \mathbf{f}$.*

f-circulant 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,

$$JZ_fJ = Z_f^T, \quad JZ_f^TJ = Z_f \tag{8}$$

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

$$Z_f^{-1} = Z_{1/f^T}. \tag{9}$$

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

Theorem 8. *(See [CPW74].) We have $Z_1(\mathbf{v}) = \Omega^{-1}D(\Omega\mathbf{v})\Omega$. More generally, for any $f \neq 0$, we have $Z_{f^n}(\mathbf{v}) = V_f^{-1}D(V_f\mathbf{v})V_f$ where $\Omega = (\omega_n^{ij})_{i,j=0}^{n-1}$ is the $n \times n$ matrix of DFT, $D(\mathbf{u}) = \operatorname{diag}(u_i)_{i=0}^{n-1}$ for a vector $\mathbf{u} = (u_i)_{i=0}^{n-1}$ and the matrix $V_f = \Omega \operatorname{diag}(f^i)_{i=0}^{n-1}$ of (4).*

The complexity of computations with Toeplitz, Hankel, Cauchy and Vandermonde matrices. Theorems 3 and 8 combined support numerically stable computation of the product by a vector of an f-circulant matrix $Z_e(\mathbf{u})$ (as well as of its inverse if the matrix is nonsingular) by using $O(n \log n)$ flops. We can extend this cost bound to multiplication of a Toeplitz matrix T of Table 1 by a vector, by embedding the matrix into a $2^k \times 2^k$ circulant matrix for $k = \lceil \log_2(2n-1) \rceil$. Each of pre- and post-multiplication by the matrix J, that is the cyclic interchange of rows or columns, transforms a Toeplitz matrix into a Hankel matrix and vice versa, and therefore transforms accordingly the algorithms for matrix inversion and solving a linear systems of equations.

Numerically unstable algorithms using nearly linear number of flops (namely $O(n \log^2 n)$ flops) are known for multiplying general Vandermonde and Cauchy $n \times n$ matrices by a vector and solving Toeplitz, Hankel, Vandermonde and Cauchy nonsingular linear systems of n equations (cf. [P01, Chapter 2 and 3]). Numerically stable known algorithms for all these problems run in quadratic arithmetic time, except that numerically stable algorithms of [MRT05], [CGS07] and [XXG12] approximate the solution of nonsingular Toeplitz linear systems in nearly linear time. We seek extension of the latter algorithms to Vandermonde and Cauchy computations.

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, the τ-rank, generators, and τ-generators of the displacement of a matrix M (for a fixed operator matrices A and B and tolerance τ) are called *displacement rank* (denoted $d_{A,B}(M)$), *τ-displacement rank*, *displacement generator*, and *τ-displacement generator* of the matrix M, respectively (cf. [KKM79]), [P01], [BM01]). In our Theorems 11 and 13 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, which we define next.

Definition 9. 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 11 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 $\mathcal{V}_{\mathbf{s}}$, $\mathcal{V}_{\mathbf{s}}^T$, 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 10. (Cf. Definition 6.) In the case where the vectors $\mathbf{s}$ and $\mathbf{t}$ turn into the vectors $e(\omega_n^{i-1})_{i=1}^n$ and $f(\omega_n^{i-1})_{i=1}^n$ for some scalars e and f, we 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 they 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.

One can readily verify the following results.

Theorem 11. 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

$$VZ_e = D_{\mathbf{s}}V - (s_i^n - e)_{i=1}^n \mathbf{e}_n^T, \tag{10}$$

$$Z_e^TV^T = V^TD_{\mathbf{s}} - \mathbf{e}_n((s_i^n - e)_{i=1}^n)^T. \tag{11}$$

Consequently the displacements $D_{\mathbf{s}}V - VZ_e$ and $Z_e^TV^T - V^TD_{\mathbf{s}}$ have rank at most 1 and vanish if

$$s_i^n = e \text{ for } i = 1, \ldots, n. \tag{12}$$

(c) For two vectors $\mathbf{s} = (s_i)_{i=1}^n$ and $\mathbf{t} = (t_i)_{i=1}^n$ having $2n$ distinct components, a Cauchy matrix C of Table 1, and the vector $\mathbf{e} = (1, \ldots, 1)^T$ of dimension n, filled with ones, we have

$$D_{\mathbf{s}}C - CD_{\mathbf{t}} = \mathbf{e}\mathbf{e}^T, \quad \text{rank}(D_{\mathbf{s}}C - CD_{\mathbf{t}}) = 1. \tag{13}$$

The following theorem shows that variation of the scalars e and f, defining the operator matrices Z_e and Z_f, makes negligible impact on the matrix structure, and so the classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, and $\mathcal{V}^T$ do not depend on the choice of these scalars.

Theorem 12. *For two scalars e and f and five matrices A, B, C, D, and M we have $d_{C,D}(M) - d_{A,B}(M) \leq 1$ 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 $B = D$, $A = Z_e^T$, $C = Z_f^T$.*

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)					

Proof. The matrix $Z_b - Z_c = (b-c)\mathbf{e}_1\mathbf{e}_n^T$ has rank at most 1 for any pair of scalars b and c. Therefore the matrices $(Z_bM - MB) - (Z_cM - MB) = Z_bM - Z_cM = (Z_b - Z_c)M$ and $(AZ - MZ_b) - (AZ - MZ_c) = -M(Z_b - Z_c)$ have ranks at most 1. $\square$

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 11 and 13. $\mathcal{V}_{\mathbf{s}}^{-1}$ 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. We obtain the pairs of their associated operator matrices by interchanging the matrices in the pairs of the operator matrices for the classes $\mathcal{V}_{\mathbf{s}}$ and $\mathcal{V}_{\mathbf{s}}^{T}$, respectively (see equation (17) of the next section).

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 11 and Table 2. See some of these and other expressions for various classes of structured matrices through their generators in [GO94], [P01, Sections 4.4 and 4.5], and [PW03].

Theorem 13. *Suppose $s_1, \dots, s_n, t_1, \dots, t_n$ are $2n$ distinct scalars, $\mathbf{s} = (s_k)_{k=1}^n$, $\mathbf{t} = (t_k)_{k=1}^n$, $V = (s_i^{k-1})_{i,k=1}^n$, $C = (\frac{1}{s_i - t_k})_{i,k=1}^n$, e and f are two distinct scalars, $\mathbf{f}_1, \dots, \mathbf{f}_d, \mathbf{g}_1, \dots, \mathbf{g}_d$ are $2d$ vectors of dimension n, $\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_n$ 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*

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

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

(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 4.4.4]. Part (c) is taken from [P01, Example 1.4.1]. To prove part (v), combine the equations $D_{\mathbf{t}}M - MZ_e = FG^T$ and $Z_eZ_{1/e}^T = I$ (cf. (9)) and deduce that $M - D_{\mathbf{t}}MZ_{1/e}^T = -F(Z_{1/e}G)^T$. Then obtain from [P01, Example 4.4.6 (part b)] that $M = e\operatorname{diag}(\frac{1}{t_i^n - e})_{i=1}^n \sum_{j=1}^n \operatorname{diag}(\mathbf{f}_j)VZ_{1/e}(Z_{1/e}\mathbf{g}_j)^T$. 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 $D_{\mathbf{t}}M^T - M^TZ_e = -GF^T$. From part (v) obtain $M^T = \operatorname{diag}(\frac{1}{e - t_i^n})_{i=1}^n \sum_{j=1}^n \operatorname{diag}(\mathbf{g}_j)VZ_e(J\mathbf{f}_j)$. Transpose this equation and arrive at part (v^T). $\square$

By combining the estimates of the previous section for the cost of multiplication by a vector of Toeplitz, Hankel, Vandermonde, transpose of Vandermonde and Cauchy matrices with Theorem 13 we obtain the following results.

Theorem 14. *Given a vector* $\mathbf{v}$ *of a dimension* n*, 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 15. By virtue of Theorem 13 the displacement operators $M \to AM - MB$ are nonsingular provided that $e \neq f$ in parts (t) and (h) and that $t_i^n \neq e$ for $i = 1, \dots, n$ in parts (v) and (v^T). We can apply Theorem 12 to satisfy these assumptions.

Remark 16. (Cf. part (ii) of Theorem 5.) Parts (v) and (c) of Theorem 13 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

We can pairwise multiply and invert structured matrices faster if we express the inputs and 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 17. *Assume five matrices* A, B, C, M *and* N *and a pair of scalars* α *and* β*. Then as long as the matrix sizes are compatible we have*

$$A(\alpha M + \beta N) - (\alpha M + \beta N)B = \alpha(AM - MB) + \beta(AN - NB), \tag{14}$$

$$A^T M^T - B^T M^T = -(BM - MA)^T, \tag{15}$$

$$A(MN) - (MN)C = (AM - MB)N + M(BN - NC). \tag{16}$$

Furthermore for a nonsingular matrix M *we have*

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

Corollary 18. *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 17 we have*

$$\begin{aligned}
F_{A,B}(\alpha M + \beta N) &= (\alpha F_{A,B}(M) \mid \beta F_{A,B}(N)), \\
G_{A,B}(\alpha M + \beta N) &= (G_{A,B}(M) \mid G_{A,B}(N)), \\
F_{A,B}(M^T) = -G_{B^T,A^T}(M^T),\ G_{A,B}(M^T) &= F_{B^T,A^T}(M^T), \\
F_{A,C}(MN) &= (F_{A,B}(M) \mid MF_{B,C}(N)), \\
G_{A,C}(MN) &= (N^T G_{A,B}(M) \mid G_{B,C}(N)), \\
F_{A,B}(M^{-1}) = -M^{-1}G_{B,A}(M),\ G_{A,B}(M^{-1}) &= M^{-1}F_{B,A}(M).
\end{aligned}$$

Consequently

$$\begin{aligned}
d_{A,B}(\alpha M + \beta N) &\leq d_{A,B}(M) + d_{A,B}(N), \\
d_{A,B}(M^T) &= d_{B^T,A^T}(M), \\
d_{A,C}(MN) &\leq d_{A,B}(M) + d_{B,C}(N), \\
d_{A,B}(M^{-1}) &= d_{B,A}(M).
\end{aligned}$$

The corollary and Theorem 13 together reduce the inversion of a nonsingular $n \times n$ matrix M given by its displacement generator of a length d to solving $2d$ linear systems of equations with this coefficient matrix M, rather than the n linear systems $M\mathbf{x}_i = \mathbf{e}_i$, $i = 1, \dots, n$ for general matrix M.

Given short displacement generators for the matrices M and N, we can apply Corollary 18 and readily express short displacement generators for the matrices M^T, $\alpha M + \beta N$, and MN through

Table 3: Operator matrices for matrix product

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

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.

Some fast inversion algorithms such as the divide-and-conquer MBA algorithm of [M80] and [BA80] involve auxiliary matrices with displacement generators whose lengths exceed the displacement rank shared by the input and output matrices. If uncontrolled, the divide-and-conquer process can blow up the length of the displacement generators and the computational cost (see [P01, Chapter 5]), but one can apply Theorem 1 to recompress the generators and obtain the following result.

Corollary 19. *[M80], [BA80]. The MBA algorithm computes a displacement generator of the inverse of a nonsingular $n \times n$ matrix M by using $O(d^2 n \log^s n)$ flops where $s = 2$ if M is from the class $\mathcal{T}$ or $\mathcal{H}$ and $s = 3$ if M is from the class $\mathcal{V}$, $\mathcal{V}^T$, or $\mathcal{C}$.*

5 Transformation of Matrix Structures

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. Thus these pairs represent the structure of the matrix classes.

Theorem 17 shows the impact of elementaty matrix operations on the associated operator matrices A and B. For linear combinations, transposes and inverses the original pair (A, B) either stays invariant or changes into $(-B^T, A^T)$ or $(-B, A)$, respectively. If the inputs are in any of the classes $\mathcal{T}$, $\mathcal{H}$, and $\mathcal{C}_{\mathbf{s},\mathbf{t}}$, then so are the outputs, whereas the transposition maps the classes $\mathcal{V}$ and $\mathcal{V}^T$ into one another and the inversion maps them into the classes $\mathcal{V}^{-1}$ and $\mathcal{V}^{-T}$, respectively. The impact of multiplication on matrix structure is quite different. As we can see from (16) and Table 3, the map $M \rightarrow PMN$ can imply 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 structure transforms given by the maps $M \rightarrow MN$, $N \rightarrow MN$, and $M \rightarrow PMN$ for appropriate multipliers P, M, and N.

Theorem 20. *It holds that*

(i) $MN \in \mathcal{T}$ if the pair of matroces (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 20 and Table 4 hold for any choice of the multipliers P and N from the indicated classes. To simplify the computation of 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 $\mathcal{C}_{\mathbf{p},\mathbf{r}}$ of the table and the theorem, 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 $f(\omega_n^{i-1})_{i=1}^n$. These multipliers are quasiunitary where $|f| = 1$. By combining Corollary 18 and Theorem 20 we obtain the following result.

Table 4: Mapping matrix structures by means of multiplication

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

Corollary 21. *Given a displacement generator of a length d for an $n \times n$ matrix M of any of the classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, $\mathcal{V}^T$, and $\mathcal{C}$, $O(dn\log^2 n)$ 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 preprocessings $M \to PMN$ with appropriate structured multipliers P and N. For an example we can decrease the complexity bound of $O(d^2 n\log^3 n)$ of Corollary 19 to $O(d^2 n\log^2 n)$ by applying canonical transformations (1) of the matrices of the bottleneck classes $\mathcal{V}$, $\mathcal{V}^T$ and $\mathcal{C}$ into the "easier" matrices of the classes $\mathcal{T}$ and $\mathcal{H}$.

5.2 The impact on displacements

In the canonical maps of Theorem 20 the displacement ranks grow by at most 2 but possibly less than that, as this is implied by Theorems 22 and 23 below. In our *constructive proofs* of these theorems we also specify the multipliers P and N and compute the displacement generators for the products PMN of some maps of Theorem 20. In some maps we set $P = I$ or $N = I$, thus omitting one of the multipliers. We show no maps where the matrices M or PMN belong to th classes $\mathcal{V}^T$, $\mathcal{V}^{-1}$, or $\mathcal{V}^{-T}$, but they can be generated from the maps with $M \in \mathcal{V}$ or $PMN \in \mathcal{V}$ by means of transposition and inversion.

Theorem 22. *Given a displacement generator of a length d for a structured matrix M of any of the four classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$, and $\mathcal{C}$, 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 $\mathcal{H}$. We yield dispacement generators of at most lengths $d+1$ in the maps $M \to PMN$ that support all other transitions among the classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$ and $\mathcal{C}$.*

Proof. We specify some maps $M \to MNP$ that suport the claims of the theorem. One can vary and combine these maps as well as the other maps of Theorem 20 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 $Z_eM - MZ_f = FG^T$. 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. (8) 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 $s_1 \dots, s_n$. Pre-multiply the displacement equation $Z_eM - MZ_f = FG^T$ by the Vandermonde matrix $V = (s_i^{j-1})_{i,j=1}^n$ to obtain $VZ_eM - (VM)Z_f = VFG^T$. Write $\mathbf{s} = (s_i)_{i=1}^n$ and substitute equation (10) to yield $D_{\mathbf{s}}(VM) - (VM)Z_f = VFG^T + (s_i^n - e)_{i=1}^n\mathbf{e}_n^TM = F_{VM}G_{VM}^T$ for $F_{VM} = (VF \mid (s_i^n - e)_{i=1}^n)$ 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 $VM \in \mathcal{V}$.

(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 $Z_e M - MZ_f^T = FG^T$. Post-multiply it by the matrix J to obtain $Z_e(MJ) - MZ_f^T J = FG^T J$. Express the term $MZ_f^T J = MJJZ_f^T J$ as MJZ_f (cf. (8)) to obtain $Z_e(MJ) - (MJ)Z_f = FG^T J = F(JG)^T$ 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) $\mathcal{V} \to \mathcal{H}$, $PMN = V^T M$. Assume $n+2$ scalars $e, f, s_1, \dots, s_n$, a matrix $M \in \mathcal{V}$, and its displacement generator given by $n \times d$ matrices F and G such that $D_{\mathbf{s}} M - MZ_f = FG^T$. Pre-multiply this equation by the transposed Vandermonde matrix $V^T = (s_j^{i-1})_{i,j=1}^n$ to obtain $V^T D_{\mathbf{s}} M - (V^T M)Z_f = V^T FG^T$ for $\mathbf{s} = (s_i)_{i=1}^n$. Apply equation (11) to express the matrix $V^T D_{\mathbf{s}}$ and obtain $Z_e^T(V^T M) - (V^T M)Z_f = V^T FG^T - \mathbf{e}_n((s_i^n - e)_{i=1}^n)^T M = F_{V^T M} G_{V^T M}^T$ for $F_{V^T M} = (V^T F \mid \mathbf{e}_n)$ and $G_{V^T M}^T = \begin{pmatrix} G^T \\ ((e - s_i^n)_{i=1}^n)^T M \end{pmatrix}$. So $d_{Z_e^T, Z_f}(V^T M) \le d_{D_{\mathbf{s}}, Z_f}(M) + 1$ and $V^T M \in \mathcal{H}$.

(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, $s_1, \dots, s_n, t_1, \dots, t_n$, a matrix $M \in \mathcal{V}$, and its displacement generator given by $n \times d$ matrices F and G. Post-multiply the equation $D_{\mathbf{s}} M - MZ_e = FG^T$ by the matrix JV^T where $V^T = (t_j^{i-1})_{i,j=1}^n$ is the transposed Vandermonde matrix, substutute $Z_e J = JZ_e^T$, and obtain $D_{\mathbf{s}}(MJV^T) - MJZ_e^T V^T = FG^T JV^T$ for $\mathbf{s} = (s_i)_{i=1}^n$. Apply equation (11) to express the matrix $Z_e^T V^T$ and obtain $D_{\mathbf{s}}(MJV^T) - (MJV^T)D_{\mathbf{t}} = FG^T JV^T - MJ\mathbf{e}_n((t_i^n - e)_{i=1}^n)^T = F_{MJV^T} G_{MJV^T}^T$ where $F_{MJV^T} = (F \mid MJ\mathbf{e}_n)$ and $G_{MJV^T}^T = \begin{pmatrix} G^T JV^T \\ ((e - t_i^n)_{i=1}^n)^T \end{pmatrix}$. So $d_{D_{\mathbf{s}}, D_{\mathbf{t}}}(MJV^T) \le d_{D_{\mathbf{s}}, Z_e}(M) + 1$ and $MJV^T \in \mathcal{C}$.

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 $D_{\mathbf{s}} M - MZ_e = FG^T$ by the matrix $V^{-1} = V_{\mathbf{t}}^{-1}$, for $\mathbf{t} = (t_i)_{i=1}^n$, to obtain $D_{\mathbf{s}} MV^{-1} - MZ_e V^{-1} = FG^T V^{-1}$. Pre- and post-multiply by V^{-1} equation (10) for $\mathbf{s}$ replaced by $\mathbf{t}$ and obtain $Z_e V^{-1} = V^{-1} D_{\mathbf{t}} - V^{-1}(t_i^n - e)_{i=1}^n \mathbf{e}_n^T V^{-1}$. Substitute the expression of $Z_e V^{-1}$ from this equation into above equation and obtain $D_{\mathbf{s}}(MV^{-1}) - (MV^{-1})D_{\mathbf{t}} = FG^T V^{-1} - V^{-1}(t_i^n - e)_{i=1}^n \mathbf{e}_n^T V^{-1} = F_{MV^{-1}} G_{MV^{-1}}^T$ for $F_{MV^{-1}} = (F \mid V^{-1}(t_i^n - e)_{i=1}^n)$ and $G_{V^{-1}M}^T = \begin{pmatrix} G^T V^{-1} \\ \mathbf{e}_n^T V^{-1} \end{pmatrix}$. So $d_{D_{\mathbf{s}}, D_{\mathbf{t}}}(VM) \le d_{D_{\mathbf{s}}, Z_e}(M) + 1$ and $MV^{-1} \in \mathcal{C}$.

(h) $\mathcal{C} \to \mathcal{V}$, $PMN = MV$. Assume $2n+1$ scalars e, $s_1, \dots, s_n, t_1, \dots, t_n$, 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 $\mathbf{s} = (s_i)_{i=1}^n$ and $\mathbf{t} = (t_i)_{i=1}^n$. Post-multiply this equation by the Vandermonde matrix $V = (t_i^{j-1})_{i,j=1}^n$ to obtain $D_{\mathbf{s}}(MV) - MD_{\mathbf{t}} V = FG^T V$. Express the matrix $D_{\mathbf{t}} V$ from matrix equation (10) and obtain $D_{\mathbf{s}}(MV) - (MV)Z_e = FG^T V + M(t_i^n - e)_{i=1}^n \mathbf{e}_n^T = F_{MV} G_{MV}^T$ where $F_{MV} = (F \mid M(t_i^n - e)_{i=1}^n)$ and $G_{MV}^T = \begin{pmatrix} G^T V \\ \mathbf{e}_n^T \end{pmatrix}$. So $d_{D_{\mathbf{s}}, Z_e}(MV) \le d_{D_{\mathbf{s}}, D_{\mathbf{t}}}(M) + 1$ and $MV \in \mathcal{C}$.

(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 23. *Assume $2n$ distinct scalars $s_1, \dots, s_n, t_1, \dots, t_n$, defining two vectors $\mathbf{s} = (s_i)_{i=1}^n$ and $\mathbf{t} = (t_j)_{j=1}^n$ and a nonsingular Cauchy matrix $C = C_{\mathbf{s},\mathbf{t}} = (\frac{1}{s_i - t_j})_{i,j=1}^n$ (cf. part (i) of Theorem 5). 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 (13) and deduce that $AMC-MCD_{\mathbf{t}} = (AM-MD_{\mathbf{s}})C+M\mathbf{e}\mathbf{e}^T$. Therefore $d_{A,D_{\mathbf{t}}}(MC) = \operatorname{rank}(AMC-MCD_{\mathbf{t}}) \le \operatorname{rank}((AM-MD_{\mathbf{s}})C)+1 = \operatorname{rank}(AM-MD_{\mathbf{s}})+1 = d_{A,D_{\mathbf{s}}}(M)+1$.

(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 (13) and deduce that $D_{\mathbf{s}}CM-CMB = C(D_{\mathbf{t}}M-MB)+\mathbf{e}\mathbf{e}^TM$. Therefore $d_{D_{\mathbf{s}},B}(CM) = \operatorname{rank}(D_{\mathbf{s}}CM-CMB) \le \operatorname{rank}(C(D_{\mathbf{t}}M-MB))+1 = \operatorname{rank}(D_{\mathbf{t}}M-MB)+1 = d_{D_{\mathbf{t}},B}(M)+1$. □

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

Multiplication by a Vandermonde multiplier $V = (s_i^{j-1})_{i,j=1}^n$ or by its transpose in Corollary 18 increases the length of a displacement generator by at most 1, but in the proof of Theorem 22 such a multiplication does not increase the length at all where $s_i^n = e$ for $i = 1,\dots,n$ and for a scalar e, employed in the operator matrices Z_e and Z_e^T of the Vandermonde displacement map (cf. (10) and (11)). This suggests choosing the vectors $\mathbf{s} = (e\omega_n^{i-1})_{i=1}^n$ and $\mathbf{t} = (f\omega_n^{i-1})_{i=1}^n$ and employing the DFT-based multipliers V_e and V_f (cf. (4)) wherever we are free to choose these vectors and multipliers. In particular we can choose such DFT-based multipliers in our maps supporting part (g) of Theorem 22, and then we would output matrices of the class $\mathcal{CV}$ having 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 7). We deduce the following results by reexamining the proof of Theorem 22 and applying transposition.

Theorem 24. *Some appropriate canonical DFT-based multipliers from the proof of Theorem 22 for the basic vectors* $\mathbf{s} = (e\omega_n^{i-1})_{i=1}^n$ *and* $\mathbf{t} = (f\omega_n^{i-1})_{i=1}^n$ *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 where* $|e| = |f| = 1$.

By combining our second map of the proof of part (g) of Theorem 22 with our map from its part (b) and choosing $\mathbf{t} = (f\omega_n^{i-1})_{i=1}^n$, we can obtain canonical DFT-based transforms $\mathcal{T} \to \mathcal{C} = \Omega\mathcal{T}\operatorname{diag}(f^{i-1})_{i=1}^n\Omega^H$, 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.

Theorem 25. *Suppose* $Z_1M - MZ_{-1} = FG^T$ *for an* $n\times n$ *matrix* M *and* $n\times d$ *matrices* F *and* G *and write* $P = \Omega_n$, $N = \omega_{2n}^{-1}\Omega_n^H$, $C = PMN$, $D = \operatorname{diag}(\omega_{2n}^{i-1})_{i=1}^n$, *and* $D = D_0^2 = \operatorname{diag}(\omega_n^{i-1})_{i=1}^n$. *Then* $DC - \omega_{2n}CD = F_CG_C^T$ *for* $F_C = \Omega_nF$ *and* $G_C = \omega_{2n}\Omega_nD_0G$.

This theorem and the supporting canonical DFT-based map $\mathcal{T} \to \mathcal{C}$ are the special cases of Theorem 22 and its transforms of matrix structures, but they appeared in [H95] as corollaries of Theorem 8. In his letter of 1991, reproduced in [P11, Appendix C], G. Heinig acknowledged studying the paper [P90], but his alternative derivation in [H95] appeared ad hoc and has defined a more narrow class of transforms of matrix structures than Theorem 22, extending [P90]. Heinig's specialization of the structure transformation method, however, has paved way to the subsequent strong demonstration of the power of the method in [GKO95], [G98], [MRT05], [R06], [CGS07], [XXG12] and has specified an efficient quasiunitary map $\mathcal{T} \to \mathcal{C}$ above, which employed the uniform distribution of the $2n$ knots $s_1,\dots,s_n,t_1,\dots,t_n$ on the unit circle $\{z:\ |z|=1\}$.

6 HSS matrices

The following class of structured matrices extends the class of banded matrices and their inverses.

Definition 26. Hereafter "*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.

This definition is one of a number of similar definitions of such matrices, also known under the names of quasiseparable, weakly, recursively or sequentially semiseparable matrices, as well as matrices with low Hankel rank and rank structured matrices. See [VVGM05], [VVM07], [VVM08], and the bibliography therein on the long history of the study of these matrix classes and see [GR87], [LRT79], [PR93] on the related subjects of Multipole and Nested Dissection algorithms.

A banded matrix B having a lower bandwidth l and an upper bandwidth u is an (l,u)-HSS matrix, and so is its inverse B^{-1} if the matrix B is nonsingular. It is well known that such a banded $n \times n$ matrix can be multiplied by a vector by using $O((l+u)n)$ flops, whereas $O((l+u)^2 n)$ flops are sufficient to solve a nonsingular linear system of n equations with such a coefficient matrix. Both properties have been extended to (l,u)-HSS matrices of the size $n \times n$ (see [MRT05], [CGS07], [XXG12]). Furthermore, like the matrices of the classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$ and $\mathcal{C}$, such an HSS matrix allows its compressed representation where one defines its generalized generator that readily expresses its n^2 entries via $O((l+u)n)$ parameters. The inverse of a nonsingular (l,u)-HSS $n \times n$ matrix M is also an (l,u)-HSS $n \times n$ matrix, and a generator expressing the inverse via $O((l+u)n)$ parameters can be computed by using $O((l+u)^2 n)$ flops. See [XXG12] and references therein on the supporting algorithms and their efficient implementation.

7 Numerical ranks of Cauchy and Vandermonde matrices

Next we bound numerical rank for a large subclass of the class of Cauchy and Cauchy-like matrices. In the next section we extend this study to approximate CV and CV-like matrices by HSS matrices.

Definition 27. A pair of complex points s and t is (θ, c)-*separated* for $\theta < 1$ and a complex point c 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). $\delta_{c,\mathbb{S}} = \min_{s \in \mathbb{S}} |s-c|$ and $\delta_{c,\mathbb{T}} = \min_{t \in \mathbb{T}} |t-c|$ denote the distances of the center c from the sets $\mathbb{S}$ and $\mathbb{T}$, respectively.

Lemma 28. *[R85]. Suppose two complex points s and t are (θ, c)-separated from one another for $0 \le \theta < 1$ and write $q = \frac{t-c}{s-c}$, $|q| \le \theta$. Then for every positive integer k we have*

$$\frac{1}{s-t} = \frac{1}{s-c} \sum_{i=0}^{k-1} \frac{(t-c)^k}{(s-c)^k} + \frac{q_k}{s-c} \text{ where } |q_k| \le \theta^k/(1-\theta). \tag{18}$$

Proof. $\frac{1}{s-t} = \frac{1}{s-c} \ \frac{1}{1-q} = \frac{1}{s-c} \sum_{i=0}^{\infty} q^i = \frac{1}{s-c} (\sum_{i=0}^{k} q^i + \sum_{i=k}^{\infty} q^i) = \frac{1}{s-c} (\sum_{i=0}^{k} q^i + \frac{q^k}{1-q})$. □

Corollary 29. *(Cf. [MRT05], [CGS07, Section 2.2].) Suppose $C = (\frac{1}{s_i - t_j})_{i,j=1}^n$ is a Cauchy matrix defined by two sets of parameters $\mathbb{S} = \{s_1, \dots, s_n\}$ and $\mathbb{T} = \{t_1, \dots, t_n\}$. Suppose these sets are* (θ, c)-separated *from one another for $0 < \theta < 1$ and a scalar c and write*

$$\delta = \delta_{c,\mathbb{S}} = \min_{i=1}^{n} |s_i - c|. \tag{19}$$

Then for every positive integer k it is sufficient to use $2kn + 4n$ ops to compute two matrices

$$F = (1/(s_i - c)^h)_{i,h=1}^{n,k+1}, \ G^T = ((t_j - c)^h)_{j,h=0}^{n,k} \tag{20}$$

that support the representation of the matrix C as $C = \widehat{C} + E$ where

$$\widehat{C} = FG^T, \ \text{rank}(\widehat{C}) \le k+1, \tag{21}$$

$$E = (e_{i,j})_{i,j=1}^n, \ |e_{i,j}| \le \frac{q^k}{(1-q)\delta} \text{ for all pairs } \{i,j\}, \tag{22}$$

and so $||E|| \le nq^k/((1-q)\delta)$.

Proof. Apply (18) for $s = s_i$, $t = t_j$ and all pairs $\{i, j\}$ to deduce (22). $\square$

Here are three immediate extensions of the theorem and the corollary.

(i) We can replace $\delta = \delta_{c,\mathbb{S}} = \min_{i=1}^{n} |s_i - c|$ by $\delta = \delta_{c,\mathbb{T}} = \min_{j=1}^{n} |t_j - c|$ because $C_{\mathbf{s},\mathbf{t}}^T = -C_{\mathbf{t},\mathbf{s}}$ (cf. (2)).

(ii) By virtue of part (c) of Theorem 13 we can extend the bounds of Corollary 29 from a Cauchy matrix $C_{\mathbf{s},\mathbf{t}}$ to a Cauchy-like matrix of the class $\mathcal{C}_{\mathbf{s},\mathbf{t}}$, given with a displacement generator (F, G) of a length d. In this extension the rank bound (21) increases by a factor of d and the error norm bound (22) increases by a factor of $d\ ||F||\ ||G||$.

(iii) Already for moderately large integers k the upper bounds of (22) are small unless the values $1 - \theta > 0$ and δ of (19) are small. Then Corollary 29 implies an upper bound $k + 1$ on the numerical rank of the large subclass of Cauchy matrices $C = (\frac{1}{s_i - t_j})_{i,j=1}^n$ whose parameter sets $\mathbb{S} = \{s_1, \ldots, s_n\}$ and $\mathbb{T} = \{t_1, \ldots, t_n\}$ are (θ, c)-separated from one another for an appropriate center c. If this property holds for two subsets of the sets $\mathbb{S}$ and $\mathbb{T}$ that define an $n \times l$ or an $l \times n$ Cauchy submatrix where $l > k + 1$, then the l rows or columns of this submatrix form a nearly rank deficient matrix, which means that the matrix C is ill conditioned. Apply a canonical DFT-based quasiunitary map $\mathcal{V} \to \mathcal{CV}$ that supports part (g) of Theorem 22 (see Theorem 24) and deduce that a Vandermonde matrix $V_{\mathbf{t}}$ is ill conditioned unless its knots from the set $\mathbb{T} = \{t_1, \ldots, t_n\}$ are close enough to all or almost all knots of the set $\{\omega_n^{i-1}\}_{i=1}^n$ of the nth roots of 1, scaled by a scalar e, $|e| = 1$. This implies (cf. [GI88]) that except for a narrow subclass all Vandermonde matrices are ill conditioned.

8 HSS approximation of CV and CV-like matrices

Theorem 30. *Assume positive integers g, h and n, a scalar e, and a Cauchy matrix $C = C_{\mathbf{s},e} = (\frac{1}{s_i - t_j})_{i,j=1}^n$ such that $t_j = e(\omega_n^{j-1})$ for $j = 1, \ldots, n$ (cf. Section 2.2), $gh = n$, n is not small, and $|e| = 1$. Then there is a permutation $n \times n$ matrix P such that CP is a $3 \times g$ block matrix with block columns $(C_{j,-}^T \mid \Sigma_j^T \mid C_{j,+}^T)^T$, $j = 0, \ldots, g-1$, where the diagonal blocks Σ_j have sizes $n_j \times h$, and the rows of the blocks Σ_j and Σ_k lie in pairwise distinct sets of rows of the matrix CP unless $|j - k| \leq 1$ or $|j - k| = g - 1$ (and so the blocks $\Sigma_1, \ldots, \Sigma_g$ together have at most $3hn$ entries), whereas every matrix $(C_{j,-}^T \mid C_{j,+}^T)^T$ is an $h \times (n - n_j)$ Cauchy matrix defined by the sets of parameters that are $(1/2, c_j)$-separated from one another for some scalars c_j lying on the unit circle $\{z : \ |z| = 1\}$ and at the distance of at least $0.5h/n^2$ from the set $\mathbb{S}_j$.*

Proof. Represent the knots $s_1, \ldots, s_n$ of the set $\mathbb{S}$ in polar coordinates, $s_i = r_i \exp(2\pi\phi_i\sqrt{-1})$ where $r_i \geq 0$, $0 \leq \phi_i < 2\pi$, $\phi_i = 0$ if $r_i = 0$, and $i = 0, 1, \ldots, n-1$. Re-enumerate all values ϕ_i to have them in nonincreasing order and to have $\phi_0^{(\text{new})} = \min_{i=0}^n \phi_i$ 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 and that $e = 1$. Let $\mathbb{S}_j = \{s_j\}_j \in \mathbb{S}$ and $\mathbb{T}_j = \{\omega_n^l\}_{l=jh}^{j(h+1)-1} \in \mathbb{T}$ denote the sets of knots lying in the semi-open sectors of the complex plane bounded by the pairs of rays from the origin to the points ω_n^{jh} and $\omega_n^{(j+1)h}$, respectively. Namely denote by $\mathbb{S}_j$ and $\mathbb{T}_j$ the subsets of the sets $\mathbb{S}$ and $\mathbb{T}$ made up of the knots whose arguments ϕ_j satisfy $2\pi jh/n \leq \phi_j < 2\pi(j(h+1) - 1)/n$, $j = 0, \ldots, g - 1$.

Write $\alpha(a, b)$ to denote the arc of the unit circle $\{z : \ |z| = 1\}$ with the end points a and b. For every j, $j = 1, \ldots, g$, choose a center c_j on the arc $\alpha(\omega_{4n}^{(4j+1)h}, \omega_{4n}^{(4j+3)h})$. This arc has the length $\pi h/n$ and shares the midpoint $\omega_{2n}^{(2j+1)h}$ with the arc $\alpha(\omega_n^{jh}, \omega_n^{(j+1)h})$, having the length $2\pi h/n$. Choose the center c_j at the distance at least $2h/n^2$ from the set $\mathbb{S}$ (as we required). This is possible because the set has exactly n elements. For $j = 0, \ldots, g-1$, index by $jh, \ldots, j(h+1) - 1$ the columns shared by the blocks $C_{j,-}$, Σ_j and $C_{j,+}$ and index the rows of the blocks Σ_j by the indices of the elements of the set $\mathbb{S}_{j-1} \cup \mathbb{S}_j \cup \mathbb{S}_{j+1}$. Note that the sets $\mathbb{S}_j$ and $\mathbb{T}_k = \{\omega_n^l\}_{(k-1)h}^{kh-1}$ are $(1/2, c_j)$-separated from one another unless $|j - k| \leq 1$ or $|j - k| = g - 1$, and this implies the separation property claimed in the theorem. $\square$

Apply Corollary 29 for $q = 1/2$, $\delta = 0.5h/n^2$, $C = (C_{u,-} \mid C_{u,+})^T$, and $u = 1, \ldots, g$ and obtain the following corollary.

Corollary 31. *The matrix PC of Theorem 30 can be represented as*

$$PC = \Sigma + \widehat{C} + E \tag{23}$$

where Σ is the block diagonal matrix $\mathrm{diag}(\Sigma_u)_{u=1}^g$, $\mathrm{rank}(\widehat{C}) \le (k+1)g$, $E = (e_{i,j})_{i,j=1}^n$, $|e_{i,j}| \le n^2 2^{2-k}/h$ for all pairs $\{i,j\}$, and so $||E|| \le n^3 2^{2-k}/h$.

Remark 32. Theorem 30 and the corollary can be immediately extended to the case where h does not divide n (in this case write $g = \lceil n/h \rceil$) as well as to the case where $C = (\frac{1}{s_i - t_j})_{i,j=1}^n$ for $s_i = e\omega^{i-1}$ for all i and $|e| = 1$ (because $C_{\mathbf{s},e} = -C_{e,\mathbf{s}}^T$ (cf. (2))). Theorem 13 implies an extension to the matrices M of the class $\mathcal{CV}$, with the increase of the rank bound by a factor of d and with the increase of the approximation norm bound by a factor of $d\ ||F||\ ||G||$ provided the matrix M is given with its displacement generator (F, G) of a length d. The proof technique of Theorem 30 enables various further extensions. Clearly one can allow any variation of the set $\mathbb{T}$ as long as its elements can be partitioned into h-tuples, each lying on or near the arc of the unit circle $\{z : |z| = 1\}$ with the endpoints ω_n^{jh} and $\omega_n^{(j+1)h}$. Furthermore the proof can be readily extended to the case where a line interval of a length between 1 and 2 (say) lying on the complex plane not very far from the origin (or on an approximation of such a line interval by a segment of a curve) replaces the unit circle $\{z : |z| = 1\}$ and where the set $\mathbb{T}$ can be partitioned into h-tuples that are more or less equally spaced on this interval (or the segment).

The block diagonal matrix Σ has at most $3hn$ entries. The matrix $\widehat{C}$ consists of the off-diagonal blocks. By combining Theorem 30 and Corollary 31 with the HSS techniques of [GR87], [MRT05], [CGS07], [XXG12], deduce that for a positive constant b and the integer $k = \lceil 3(b+2)\log_2 n \rceil$, the matrix $\widehat{C}$ of (23) is an (l,u)-HSS matrix where $l + u \le ckh$, $h \le c' \log n$, $n^3 2^{2-k}/h \le 2^{-b}$, and c and c' are two constants.

9 Multiplication of the matrices of the classes $\mathcal{CV}$, $\mathcal{V}$, $\mathcal{V}^T$, and $\mathcal{C}$ and their inverses by vectors

Suppose $\mu(M)$ denotes the minimum number of flops sufficient for multiplying a matrix M by a vector and estimate $\mu(C) = \mu(PC)$ for the matrices of Corollary 31. The matrix Σ has at most $3hn$ nonzero entries, and so $\mu(\Sigma) \le 6hn - n$. Furthermore $\mu(\widehat{C}) = O(n \log n)$ because the matrix $\widehat{C}$ has the (l,u)-HSS structure for $l + u \le ckh$ and $h \le c' \log n$ (see Section 6). Let us summarize the estimates for the CV matrices with an extension to the matrices of the classes $\mathcal{CV}$, $\mathcal{V}$, and $\mathcal{V}^T$.

Theorem 33. *(See Remark 36.) Assume a positive scalar b, a complex e such that $|e| = 1$, and two vectors $\mathbf{f}$ and $\mathbf{s}$ of dimension n. (i) Then one can approximate the product $M\mathbf{f}$ within the error norm bound $2^{-b}\ ||M||\ ||\mathbf{f}||$ by using $O(bn \log n)$ flops provided that M is a CV, Vandermonde or transposed Vandermonde $n \times n$ matrix $C_{\mathbf{s},e}$, $C_{e,\mathbf{s}}$, $V_{\mathbf{s}}$ or $V_{\mathbf{s}}^T$, respectively. (ii) The flop bound for solving a nonsingular linear system of n equations with the coefficient matrix in the above classes increases versus part (i) by a factor of $\log n$ and the error norm bounds increases by a factor of $||M^{-1}||/||M||$. (iii) The flop bounds of parts (i) and (ii) also hold for approximate evaluation of a polynomial of degree $n-1$ at n points and for approximate interpolation to this polynomial from its n values, respectively. (iv) The flop bounds of parts (i) and (ii) increase by a factor of d, whereas the error norm bounds increase by a factor of $d\ ||F||\ ||G||$ where M is a matrix from the class $\mathcal{C}_{\mathbf{s},e}$, $\mathcal{C}_{e,\mathbf{s}}$, $\mathcal{V}_{\mathbf{s}}$ or $\mathcal{V}_{\mathbf{s}}^T$ (having the structure of CV, Vandermonde or transposed Vandermonde type) given with a displacement generator (F, G) of a length d.*

Proof. Summarize our estimates above to deduce the bound of part (i) in the case of CV matrices $C_{\mathbf{s},e}$ and $C_{e,\mathbf{s}}$. Apply Theorem 2 to estimate the approximation errors of solving the linear systems of equations and extend the bounds of part (i) to part (ii). To extend the estimates of parts (i) and (ii) to the case of Vandermonde matrices $V_{\mathbf{s}}$, apply the canonical DFT-based specialization of a map supporting part (g) of Theorem 22 for the DTF-based matrix $V = V_f$ where $|f| = 1$, and so $||V_f|| = \sqrt{n}$. The map increases the approximation error norm (versus the case of CV matrices $C_{\mathbf{s},e}$ and $C_{e,\mathbf{s}}$) by a factor of $\sqrt{n}\min_{i=1}^{n} \frac{1}{|s_i^n - f|}$. Choose a complex f, $|f| = 1$, that keeps this factor below $3n\sqrt{n}$. Compensate for this increase of the norm bound by adding $\log_2(3n\sqrt{n})$ to the value k. Similarly multiply a transposed Vandermonde matrix by a vector, transpose a map that supports part (g) of Theorem 22, and employ equation (2). Extend the results of parts (i) and (ii) to part (iii) by applying Theorem 4. Extend them to part (iv) by applying parts (v), (v^T), and (c) Theorem 13, choosing a scalar e in parts (v) and (v^T) such that $\min_{i=1}^{n} |e - s_i^n| \geq 1/2$ (say), and increasing the integer parameter k by $\lceil \log(en) \rceil$ (to compensate for the exceess of the norms $||Z_e(\mathbf{f}_i)||$ and $||Z_e(\mathbf{g}_i)||$ above $||F||$ and $||G||$, respectively). □

Remark 34. One can ignore the HSS structure of the matrix $\widehat{C}$ and still approximate the matrix product $C_{\mathbf{s},e}\mathbf{f}$ at the cost bounds that are smaller than the known bounds by a factor of $\sqrt{n/\log n}$. Indeed choose h of about $\sqrt{n \log n}$ and choose g of about $\sqrt{n/\log n}$ in Corollary 31 and obtain the matrix $\widehat{C}$ of a rank of order $\sqrt{n \log n}$. We can multiply this matrix by a vector by using $O(n\sqrt{n \log n})$ flops. The estimate is extended to the overall cost of multiplying the matrix $\Sigma + \widehat{C}$ by a vector because we can multiply the matrix Σ by a vector by using $6hn - n$ flops and because $h = O(\sqrt{n \log n})$.

Remark 35. We can extend Theorem 33 similarly to the extensions of Theorem 30 and Corollary 31 in the second part of Remark 32.

Remark 36. The algorithms supporting Theorem 33 can be naturally partitioned into two stages. At first we apply canonical DFT-based transformations of Theorems 22–24 and 33 and Corollary 29 to reduce our tasks to computations with HSS matrices. At this stage we propose a novel specialization of the approach of [P90]. Then it remains to apply the Multipole algorithms, which is both powerful and well developed. We perform the former (FFT-based) stage by applying $O(n \log n)$ flops. The latter (Multipole/HSS) stage involves $O((l+u)n)$ flops for multiplication of an $n \times n$ HSS matrix by a vector and $O((l+u)^2 n)$ flops for solving a nonsingular HSS linear system of n equations, and we have the bound $l + u = O(\log n)$ in our case. Empirically, however, in the extensive tests in [XXG12] for HSS computations similar to ours, the value $(l + u)^2$ grew much slower than $\log n$ as n grew large, and so we can expect that the computational cost at the first (FFT) stage of the algorithms actually dominates their overall computational cost.

Clearly, the algorithms supporting Theorem 33 are efficient not only for CV and CV-like matrices, but for a larger subclass of the class of Cauchy-like matrices (cf. Remark 32). The extension to the general Cauchy and Cauchy-like matrices can lead to numerical problems, however. Here are some sketchy comments. 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. Then for a large class of vectors $\mathbf{s}$ and $\mathbf{t}$, one can extend Theorem 30 and reduce the approximation of the vectors $M\mathbf{u}$ and of the solution $\mathbf{x}$ to a linear system of n equations $M\mathbf{x} = \mathbf{u}$ (if it is nonsingular) to HSS computations (cf. Remark 32).

Furthermore for all input vectors $\mathbf{s}$, $\mathbf{t}$ and $\mathbf{u}$, we can apply our techniques of transforming matrix structures to reduce the solution $\mathbf{x}$ of the linear system $M\mathbf{x} = \mathbf{u}$ to some computations with CV matrices and to the computation of the product of the matrix M by the vector $\mathbf{e}$ as follows. Fix a scalar e, write $P = MC_{\mathbf{t},e}$ and $\mathbf{x} = C_{\mathbf{t},e}\mathbf{y}$, and note that $P\mathbf{y} = \mathbf{u}$, whereas $P \in \mathcal{C}_{\mathbf{s},e}$ is a CV matrix with the displacement generator (F_P, G_P) of length at most $d+1$ where $F_P = (F \mid M\mathbf{e})$ and $G_P = (C_{\mathbf{t},e}^T G \mid \mathbf{e})$. 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 matrices. In both cases application of the algorithms would require additional error analysis. E.g., the approximation errors of computing the matrix P would magnify the approximation errors for the vectors $\mathbf{y}$ (cf. Theorem 2) and $\mathbf{x}$.

Next we consider another extension of our techniques and make further comments on error propagation. Part (c) of Theorem 13 enables us to reduce 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 $\mathbf{v}_i = \mathrm{diag}(\mathbf{g}_i)_{i=1}^d \mathbf{u}$, $\mathbf{g}_i = G\mathbf{e}_i$, and $i = 1,\dots,d$, and to $O(n)$ additional flops, provided the matrix $M \in \mathcal{C}_{\mathbf{s},\mathbf{t}}$ is given with its displacement generator (F,G) of a length d. (For $d=1$ and $(F,G)=(1,1)$ we arrive at the problems of rational multipoint evaluation and interpolation (see part (ii) of Theorem 4).) Equation (3) reduces multiplication $C_{\mathbf{s},\mathbf{t}}\mathbf{v}$ to multiplication of each of the matrices $V_{\mathbf{s}}$ and $V_{\mathbf{t}}^{-1}$ by $d+1$ vectors, to one multiplication of the matrix $V_{\mathbf{t}}$ by a vector, and to $O(n)$ additional flops (cf. Theorem 4). We can apply the new fast algorithms to approximate the $2d+3$ matrix-by-vector products above, but the approximation errors can readily propagate in this application of the algorithms.

10 Conclusions

At first we revisited our approach of [P90] to the transformation of matrix structures, covered it comprehensively, and simplified its presentation by employing the Sylvester (rather than Stein) displacements and the techniques for operating with them from [P00] and [P01, Section 1.5]. Then we singled out a large subclass of Cauchy-like matrices, which we call the CV-like matrices. We closely approximated these matrices by HSS matrices and then applied the Multipole method to the latter HSS matrices. This yielded dramatic acceleration of the known numerical algorithms that approximated the products of CV and CV-like matrices by vectors and the solution of nonsingular linear systems of equations with CV and CV-like coefficient matrices. Namely the running time of the new algorithms is nearly linear, versus quadratic time required by the known algorithms. By properly transforming matrix structures we have readily extended such an acceleration of the known algorithms to the matrices having structures of Vandermonde and transposed Vandermonde types, and consequently to numerical multipoint evaluation and interpolation of polynomials.

Potential extensions and specializations include computations with confluent Vandermonde matrices, Loewner matrices, and various problems of rational interpolation such as the Nevanlinna–Pick and matrix Nehari problems (cf. [P01, Chapter 3] and the bibliography therein), where, however, the progress can be limited to the case of sufficiently well conditioned inputs. Our demonstration of the power of the transformation of matrix structures should motivate research efforts for finding new inexpensive transforms of matrix structures and their new algorithmic applications. Natural topics of further study should include the following issues:

(i) extension of our our approach to a larger class of Cauchy and Cauchy-like matrices (cf. Remark 32),

(ii) the impact of the conditioning of the input on the output errors and the running time,

(iii) the estimation of the treshold input sizes for which the proposed algorithms running in nearly linear time outperform their variant of Remark 34 and the known algorithms, running in quadratic time, and

(iv) implementation of the proposed algorithms.

The implementation should be mostly reduced to the application of the Multipole algorithms and should extend [XXG12], but the actual work should prompt the refinements toward decreasing the treshold values of part (iii).

Acknowledgements: Our research has been supported by NSF Grant CCF–1116736 and PSC CUNY Awards 64512–0042 and 65792–0043.

References

[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.

[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.

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

[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.

[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 Matrcies, *Numerische Mathematik*, **52**, 241–250, 1988.

[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.

[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.

[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.

[P93] V. Y. Pan, Decreasing the Displacement Rank of a Matrix, *SIAM Journal on Matrix Analysis and Applications*, **14**, **1**, 118–121, 1993.

[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.

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

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

[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.

[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.

[XXCB] J. Xia, Y. Xi, 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.