2013

# TR-2013003: Polynomial Evaluation and Interpolation: Fast and Stable Approximate Solution

Victor Y. Pan

# Polynomial Evaluation and Interpolation: Fast and Stable Approximate Solution

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

**Abstract**

Multipoint polynomial evaluation and interpolation are fundamental for modern algebraic and numerical computing. The known algorithms solve both problems over any field by using $O(N \log^2 N)$ arithmetic operations for the input of size $N$, but the cost grows to quadratic for numerical solution. Our study results in numerically stable algorithms that use $O(uN \log N)$ arithmetic time for approximate evaluation (within the relative output error norm $2^{-u}$) and $O(uN \log^2 N)$ time for approximate interpolation. The problems are equivalent to multiplication of an $n \times n$ Vandermonde matrix by a vector and the solution of a nonsingular Vandermonde linear systems of $n$ equations, respectively. The algorithms and complexity estimates can be also applied where the transposed Vandermonde matrices replace Vandermonde matrices. Our advance is due to employing and extending our earlier method of the transformation of matrix structures, which enables application of the HSS–Multipole method to our tasks and further extension of the algorithms to more general classes of structured matrices.

**Key words:** Polynomials, Multipoint evaluation, Interpolation, Vandermonde matrices, Cauchy matrices, Structured matrices, Transforms of matrix structures, HSS matrices, Multipole method

## 1 Introduction

Multipoint polynomial evaluation and interpolation are fundamental for modern algebraic and numerical computing. For the input defined by $N$ parameters, the classical solution algorithms use order of $N^2$ arithmetic operations. The fast algorithms (sometimes called superfast) use $O(N \log^2 N)$ arithmetic operations over any field [BP94], [P01], [GG03], but are numerically unstable. Quadratic time algorithms are still the user's choice for numerical computations, in spite of some research advances in [PSLT93], [P95], [PZHY97].

Our present study produces numerically stable algorithms for approximate multipoint polynomial evaluation and interpolation within the relative output error norm $2^{-u}$ by using $O(uN \log N)$ and $O(uN \log^2 N)$ arithmetic time, respectively. According to some empirical evidence even these nearly optimal bounds can actually be overly pessimistic (see the end of Section 7).

We have obtained our results as corollaries from similar results on the approximation (by using nearly linear arithmetic time) of the product of a Vandermonde matrix by a vector and of the solution of a nonsingular Vandermonde linear system of equations. Our techniques imply the same results where transposed Vandermonde matrices replace Vandermonde matrices.

Our progress relies on reducing the evaluation and interpolation tasks to computations with structured matrices, transformation of matrix structures, and application of the Multipole algorithms. The latter technique has previously succeeded in the paper [MRT05] (see also [CGS07], [XXG12]), for producing a numerically stable approximate solution of nonsingular Toeplitz linear systems of equations by using nearly linear arithmetic time, but we apply a distinct transformation of matrix structure and also apply new techniques that enable incorporation of the Multipole algorithms into approximate computations with a broader class of matrices, covering the class of Vandermonde matrices, their transposes and inverses.

The papers [MRT05], [CGS07], [XXG12] as well as [H95], [GKO95] and [G98] employ the same basic transformation of matrix structures. Like our transforms in this paper, it is a specialization of the general approach proposed in [P90], and we hope that our current work will renew interest to this approach.

Our present advance has potential impact to the practice of univariate polynomial root-finding. The current best package of subroutines MPSolve 2012 (second release) relies on Ehrlich–Aberth iterations, which amount essentially to recursive multipoint polynomial evaluation. MPSolve performs it in quadratic time by means of Horner's algorithm [BF00], that is slower by order of magnitude than this can be done now, based on our current progress.

We organize our paper as follows. We recall some definitions in the next section, recall the evaluation and interpolation tasks (including their matrix versions) and their currrent solution cost in Section 3, and specify fast FFT-based reduction of these tasks to each other in Section 4. We approximate Cauchy matrices of a large class by low-rank matrices in Section 5. In Section 6 we employ these approximations to accelerate by a factor of $\sqrt{N/\log N}$ the known numerical algorithms for multiplication of these matrices by a vector. In Section 7 we apply the HSS techniques to strengthen this acceleration by another factor of $\sqrt{N}$ and to yield fast numerical approximation of the product of a Vandermonde matrix by a vector and the solution of a nonsingular linear system of equations with such a matrix. In Section 8 we extend the Cauchy matrix structure by using displacement operators. We conclude the paper with short Section 9.

We refer the reader to [Pa] on an extension of this work to various computations with structured matrices and to a more comprehensive presentation of methods of the transformation of their structures.

## 2 Definitions and auxiliary results

Hereafter "op" stands for "arithmetic (field) operation".

$M = (m_{i,j})_{i,j=1}^{m,n}$ is an $m \times n$ matrix. $M^T$ and $M^H$ are its transpose and Hermitian transpose, respectively.

$(B_1 \mid \dots \mid B_k)^T$ is a $k \times 1$ block matrix with blocks $B_1, \dots, B_k$. $\mathrm{diag}(B_1, \dots, B_k) = \mathrm{diag}(B_j)_{j=1}^k$ is a $k \times k$ block diagonal matrix with diagonal blocks $B_1, \dots, B_k$.

For $1 \times 1$ blocks $b_j = B_j$, $j = 1, \dots, k$, these are a vector $\mathbf{b} = (b_j)_{j=1}^k$ of dimension $k$ and a $k \times k$ diagonal matrix $D_{\mathbf{b}} = \mathrm{diag}(b_j)_{j=1}^k$, respectively.

$I = I_n = (\mathbf{e}_1 \mid \dots \mid \mathbf{e}_n) = \mathrm{diag}(1)_{j=1}^n$ is the $n \times n$ identity matrix. Its columns $\mathbf{e}_1, \dots, \mathbf{e}_n$ are the $n$ coordinate vectors. $J = J_n = (\mathbf{e}_n \mid \dots \mid \mathbf{e}_1)$ is the $n \times n$ reflection matrix.

$I$ and $J$ are the simplest examples of Toeplitz matrices $T = (t_{i-j})_{i,j=1}^n$ and Hankel matrices $H = (h_{i+j})_{i,j=0}^{n-1}$, which are two most popular classes of structured matrices. $JH$ and $HJ$ are Toeplitz matrices if $H$ is a Hankel matrix, and vice versa. We recall the following well known result.

**Fact 2.1.** *One can multiply an $n \times n$ Toeplitz or Hankel matrix by a vector by using $O(n \log n)$ ops.*

**Fact 2.2.** *It is sufficient to perform $O(n \log^2 n)$ ops to approximate closely the solution of a nonsingular $n \times n$ Toeplitz or Hankel linear system of $n$ equations.*

*Proof.* Apply the multiplier $J$ to reduce a Hankel linear system to Toeplitz linear system and use [XXG12] to compute their approximate solutions. □

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

$M$ is a $m \times n$ unitary matrix if $M^H M = I_n$ or $MM^H = I_m$.

$\kappa(M) = ||M||\, ||M^{-1}|| \geq 1$ is the condition number of a nonsingular matrix $M$ (for a fixed matrix norm $||\cdot||$). The unitary matrices have full rank and the minimum condition number 1. A matrix is *ill conditioned* if its condition number is large (in context). $\Delta$-rank of a matrix $M$ for a positive $\Delta$ is the integer $\min_{||\tilde{M}-M||\leq\Delta}\{\text{rank}\,\tilde{M}\}$. The numerical rank is the $\Delta$-rank for a small positive $\Delta$.

**Theorem 2.1.** *(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 $\theta \leq 1/3$.*

# 3 Polynomial and rational evaluation and interpolation. Links to Vandermonde and Cauchy matrices and complexity

As we state below, multipoint polynomial and rational evaluation and interpolation problems are equivalent to multiplying Vandermonde and Cauchy matrices by vectors and solving linear systems of equations with these matrices, respectively. We refer the reader to [P01, Ch. 3] and the bibliography therein on these and other links between the problems of polynomial and rational evaluation and interpolation and structured matrices.

**Problem 1. Multipoint polynomial evaluation and Vandermonde-by-vector product.**

INPUT: $2n$ scalars $p_0, \ldots, p_{n-1}; x_1, \ldots, x_n$.

OUTPUT: $n$ scalars $v_1, \ldots, v_n$ satisfying

$$v_i = p(x_i) \text{ for } p(x) = p_0 + p_1 x + \cdots + p_{n-1}x^{n-1} \text{ and } i = 1, \ldots, n \tag{3.1}$$

or equivalently

$$V\mathbf{p} = \mathbf{v} \text{ for } V = V_{\mathbf{x}} = (x_i^{j-1})_{i,j=1}^n,\ \mathbf{p} = (p_j)_{j=0}^{n-1},\ \text{and } \mathbf{v} = (v_i)_{i=1}^n. \tag{3.2}$$

**Problem 2. Polynomial interpolation and Vandermonde linear system solving.**

INPUT: $2n$ scalars $v_1, \ldots, v_n; x_1, \ldots, x_n$, the last $n$ of them distinct.

OUTPUT: $n$ scalars $p_0, \ldots, p_{n-1}$ satisfying equations (3.1) and (3.2).

**Problem 3. Multipoint rational evaluation and Cauchy-by-vector product.**

INPUT: $3n$ scalars $s_1, \ldots, s_n; t_1, \ldots, t_n; v_1, \ldots, v_n$.

OUTPUT: $n$ scalars $v_1, \ldots, v_n$ satisfying

$$v_i = \sum_{j=1}^{n} \frac{u_j}{s_i - t_j} \text{for } i = 1, \ldots, n \tag{3.3}$$

or equivalently

$$C\mathbf{u} = \mathbf{v} \text{ for } C = C_{\mathbf{s},\mathbf{t}} = \left(\frac{1}{s_i - t_j}\right)_{i,j=1}^n,\ \mathbf{u} = (u_j)_{j=1}^n,\ \text{and } \mathbf{u} = (u_i)_{i=1}^n. \tag{3.4}$$

**Problem 4. Rational interpolation and Cauchy linear system solving.**

INPUT: $3n$ scalars $s_1, \ldots, s_n; t_1, \ldots, t_n; v_1, \ldots, v_n$, the first $2n$ of them distinct.

OUTPUT: $n$ scalars $u_1, \ldots, u_n$ satisfying (3.3) and (3.4).

Note that every rational function $r(x) = \frac{p(x)}{t(x)}$, for $p(x)$ of equation (3.1), $t(x) = \prod_{j=1}^{n}(x - t_j)$ and distinct scalars $t_1, \ldots, t_n$, can be represented as $r(x) = \sum_{j=1}^{n} \frac{u_j}{x-t_j}$, which turns into equations (3.3) if we write $v_i = r(s_i)$ for $i = 1, \ldots, n$.

The matrices $V = V_{\mathbf{x}}$ in (3.2) and $C = C_{\mathbf{s},\mathbf{t}}$ in (3.4) are called *Vandermonde* and *Cauchy* matrices, respectively. They make up two other classes of most popular structured matrices, besides the classes of Toeplitz and Hankel matrices. These matrix classes have quite distinct features. E.g., 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). Nevertheless the four classes can be linked to each other by means of structured matrix multiplication. In addition to the cited Toeplitz–Hankel link via the multiplier $J$, we have various links via Vandermonde and transposed Vandermonde multipliers (see [P90], [P01, Sections 4.7 and 4.8]). The following fact is a simple example.

**Fact 3.1.** *(i)* $H = V^T V = (\sum_{k=1}^{n} x_k^{i+j-2})_{i,j=1}^{n}$ *is a Hankel matrix for any* $m \times n$ *Vandermonde matrix* $V = (x_i^{j-1})_{i,j=1}^{m,n}$. *(ii) Given the entries* $x_1, \ldots, x_n$ *of the* $n \times n$ *matrix* $V$, *it is sufficient to perform* $O(n \log n)$ *ops to compute all entries of the matrix* $H$.

*Proof.* Verify part (i) by inspection. Support part (ii) by the algorithm of [BP94, page 34]. □

It is well known that $\det V_{\mathbf{x}} = \prod_{i<k}(x_i - x_k)$ and $\det C_{\mathbf{s},\mathbf{t}} = \prod_{i<j}(s_i - s_j)(t_i - t_j)/\prod_{i,j}(s_i - t_j)$, and so the matrices $V$ and $C$ are nonsingular (and even strongly nonsingular, that is nonsingular together with all their leading blocks) where the knots $x_1, \ldots, x_n$ and $s_1, \ldots, s_n; t_1, \ldots, t_n$, respectively, are distinct, and then Problems 2 and 4 have unique solutions.

Let us recall the complexity of the known algorithms for Problems 1–4. Horner's algorithm of 1819 evaluates the polynomial $p(x)$ of (3.1) at a single knot $x$ by using $2n-2$ ops, and this is optimal [P66]. (The algorithm was used by Newton in 1669, by a number of medieval mathematicians, and in the Nine Chapters on the Mathematical Art at the time of the Han Dynasty in China (202 BC–220 AD).) For Problem 1 with $n$ knots, $n$ applications of Horner's algorithm involve $2(n-1)n$ ops, but this is not optimal anymore. The algorithms of [F72], [F72], and [MB72] (cf. [BP94], [P01], [GG03]) solve both Problems 1 and 2 by using $O(n \log^2 n)$ ops over any field of constants, which is within a logarithmic factor from the optimum [S73], [B-O83]. For numerical computations with rounding, however, these fast algorithms are unstable, and the users choose quadratic time algorithms for Problems 1–4 (cf. [BF00], [P64], [BP70], [BEGO08]).

# 4 DFT, IDFT, FFT, IFFT, and the reduction of polynomial to rational computations

Suppose $x_i = \omega^i$ for $i = 0, \ldots, n-1$, $\omega = \exp(2\pi\sqrt{-1}/n)$ is a primitive $n$th root of 1, $V = \sqrt{n}\Omega = (\omega^{(i-1)(j-1)})_{i,j=1}^{n}$, $V^{-1}/\sqrt{n} = \Omega^H = \Omega^{-1} = \frac{1}{\sqrt{n}}(\omega^{-(i-1)(j-1)})_{i,j=1}^{n}$, $\Omega$ is a unitary matrix, $\Omega^H \Omega = I_n$. Then Problems 1 and 2 turn into the computational problems of the forward and inverse discrete Fourier transforms (hereafter *DFT* and *IDFT*). The *FFT* (Fast Fourier transform) and Inverse FFT (*IFFT*) are numerically stable algorithms that solve these problems by using $1.5n \log_2 n$ and $1.5n \log_2 n + n$ ops, respectively, if $n$ is a power of 2 (cf. [BP94, Sections 1.2 and 3.4]). Generalized FFT uses $O(n \log n)$ ops to solve these tasks for any $n$ [P01, Problem 2.4.2].

Next we reduce Problems 1 and 2 of polynomial computations to Problems 3 and 4 of rational computations with the knots $t_j = a\omega^{j-1}$, for $j = 1, \ldots, n$ and a scalar $a \neq 0$, and vice versa by using $O(n \log n)$ ops (see an alternative way in Section 8).

For a polynomial $p(x)$ of a degree $n-1$ and a scalar $a$ define the following expression,

$$\frac{p(x)}{x^n - a^n} = \sum_{j=1}^{n} \frac{u_j}{x - a\omega^{j-1}} \tag{4.1}$$

where the Lagrange interpolation formula for the knots $a, a\omega, \ldots, a\omega^{n-1}$ implies that

$$u_j = p(a\omega^{j-1})/(na^{j-1}\omega^{(j-1)(n-1)}),\ j = 1, \ldots, n. \tag{4.2}$$

Substitute $x = s_i$ into (4.1) and obtain

$$p(s_i) = (s_i^n - a^n) \sum_{j=1}^{n} \frac{u_j}{s_i - a\omega^{j-1}}, \ i = 1, \ldots, n. \tag{4.3}$$

Equations (4.2) and (4.3) enable us to reduce Problems 1 and 3 (with the knots $t_j = \omega^{j-1}$, $j = 1, \ldots, n$) to one another by means of applying DFT, computing the values $s_1^{n-1}, \ldots, s_n^{n-1}$ (both stages involve $O(n \log n)$ ops), and performing $O(n)$ additional ops.

Let us similarly apply $O(n \log n)$ ops to reduce to one another Problems 2 and 4 (with the knots $t_j = \omega^{j-1}$, $j = 1, \ldots, n$) where $n = 2^k$. Given the knots $x_i = s_i$ and the values $p(s_i)$ for $i = 1, \ldots, n$, we evaluate the coefficients $p_0, \ldots, p_{n-1}$ of $p(x)$, thus solving Problem 2, by successively performing the following stages.

**Algorithm 4.1.** *Reduction of Problem 2 to Problem 4 for roots of 1 as t-knots.*

*1. Compute the values $v_i = p(s_i)/(s_i^n - a^n)$, $i = 1, \ldots, n$ by using $O(n \log n)$ ops.*

*2. Compute the values $u_1, \ldots, u_n$ satisfying (4.3) (this is Problem 4 for $t_j = a\omega^{j-1}$, $j = 1, \ldots, n$).*

*3. Sum the regular fractions $\frac{u_j}{x - a\omega^{j-1}}$ for $j = 1, \ldots, n$ and output the coefficients $p_0, \ldots, p_{n-1}$ of the numerator polynomial of the output fraction.*

Throughout the summation at Stage 3 ensure that the denominators of all computed fractions should be binomials. To achieve this, always sum pairs of fractions having the denominators of the form $x^{2^h} - a^{2^h}\omega^{g(h)}$ and $x^{2^h} + a^{2^h}\omega^{g(h)}$. Then the denominator of the sum is also a binomial, $x^{2^{h+1}} - a^{2^{h+1}}\omega^{2g(h)}$. In this process Stage 3 has $k = \log_2 n$ levels of summation for $h = 0, 1, \ldots, k-1$, and we use $n$ multiplications and $n/2$ additions at each level, that is $1.5kn = 1.5n \log_2 n$ ops overall. This proves the desired reduction of Problem 2 to Problem 4 (with the knots $t_j = \omega^{j-1}$, $j = 1, \ldots, n$) at the cost performing $O(n \log n)$ ops where $n = 2^k$.

**Algorithm 4.2.** *Reduction of Problem 4 to Problem 2 for roots of 1 as t-knots.*
*Given the values $v_i = \sum_{j=1}^{n} \frac{p(s_i)}{s_i^n - a^n} = \frac{u_j}{s_i - a\omega^{j-1}}$ solve Problems 4 as follows. First compute the values $p(s_i) = (s_i^n - a^n)v_i$, $i = 1, \ldots, n$, by using $O(n \log n)$ ops, then compute the coefficients of the polynomial $p(x)$ (by solving Problem 2) and the values $p(a\omega^{j-1})$ (with the knots $t_j = \omega^{j-1}$, for $j = 1, \ldots, n$) by using $O(n \log n)$ ops, and finally apply exressions (4.2) to compute the values $u_1, \ldots, u_n$.*

To extend our reduction of Problems 2 and 4 to one another to the case of any positive integer $n$, apply generalized FFT. Then the cost bound of $O(n \log n)$ ops is still supported, although the overhead constant increases a little. For the general set of knots $t_1, \ldots, t_n$ the algorithms can be extended, but the cost bound would grow to $O(n \log^2 n)$ ops.

To support alternative reductions among Problems 1–4, we can employ the following matrix equation (see, e.g., [P01, Section 3.6]),

$$C_{\mathbf{s},\mathbf{t}} = \operatorname{diag}(t(s_i)^{-1})_{i=1}^{n} V_{\mathbf{s}} V_{\mathbf{t}}^{-1} \operatorname{diag}(t'(t_i))_{i=1}^{n} \tag{4.4}$$

where $\mathbf{s} = (s_i)_{i=1}^n$, $\mathbf{t} = (t_i)_{i=1}^n$, and $t(x) = \prod_{i=0}^{n-1}(x - t_i)$. and write $\mathbf{t} = (\omega^{j-1})_{j=1}^n$ to obtain $V_{\mathbf{t}} = \sqrt{n}\Omega$ and $V_{\mathbf{t}}^{-1} = \sqrt{n}\Omega^H$.

# 5 Low-rank approximation of Cauchy matrices

An $m \times n$ matrix $M$ of a rank $l$ can be nonuniquely expressed as $M = GH^T$ where the pair $\{G, H\}$ of matrices of sizes $m \times l$ and $n \times l$, respectively, is called a *generator of length $l$* for the matrix $M$, $l \geq \operatorname{rank}(M)$. The numerical (or approximate) rank of a matrix is the minimum rank of a nearby matrix, under a fixed tolerance to the approximation error norm. Next we will compute short generators of low-rank approximations for a large class of $n \times n$ Cauchy matrices.

**Definition 5.1.** *Two complex points $s$ and $t$ are $(\theta, c)$-*separated *from one another if $|\frac{t-c}{s-c}| \le \theta$ (for a* separation factor *$\theta < 1$ and a complex* separation center *$c$). Two sets of complex numbers $\mathbb{S}$ and $\mathbb{T}$ are $(\theta, c)$-*separated *from one another if every two points $s \in \mathbb{S}$ and $t \in \mathbb{T}$ are $(\theta, c)$-separated from one another. $\delta_{c,\mathbb{S}} = \min_{s\in\mathbb{S}} |s-c|$ and $\delta_{c,\mathbb{T}} = \min_{t\in\mathbb{T}} |t-c|$ denote the distances from the center $c$ to the sets $\mathbb{S}$ and $\mathbb{T}$, respectively.*

**Lemma 5.1.** *Suppose two complex points $s$ and $t$ are $(\theta, c)$-separated from one another for a positive $\theta < 1$ and a complex $c$ and write $q = \frac{t-c}{s-c}$, $|q| \le \theta$. Then for every positive integer $k$ we have*

$$\frac{1}{s-t} = \frac{1}{s-c}\sum_{i=0}^{k-1}\frac{(t-c)^k}{(s-c)^k} + \frac{q_k}{s-c} \text{ where } |q_k| = \frac{|q|^k}{1-|a|} \le \frac{\theta^k}{1-\theta}. \tag{5.1}$$

*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 5.1.** *(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 $(\theta, c)$-*separated *from one another for $0 < \theta < 1$ and a scalar $c$ and write $\delta = \delta_{c,\mathbb{S}} = \min_{i=1}^n |s_i - c|$. Then for every positive integer $k$ it is sufficient to use $2kn + 4n$ ops to compute the matrices*

$$G = (1/(s_i-c)^g)_{i,g=1}^{n,k+1},\ H^T = ((t_j-c)^h)_{j,h=0}^{n,k}, \tag{5.2}$$

*supporting the representation $C = \widehat{C} + E$ where*

$$\widehat{C} = GH^T,\ \operatorname{rank}(\widehat{C}) \le k+1, \tag{5.3}$$

$$E = (e_{i,j})_{i,j=1}^n,\ |e_{i,j}| \le \frac{q^k}{(1-q)\delta} \text{ for all pairs } \{i,j\}, \tag{5.4}$$

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

*Proof.* Apply (5.1) for $s = s_i$, $t = t_j$ and all pairs $\{i, j\}$ to deduce (5.4). □

Corollary 5.1 bounds the numerical rank of the large subclass of Cauchy matrices $C = (\frac{1}{s_i-t_j})_{i,j=1}^n$ whose parameter sets $\mathbb{S} = \{s_1, \dots, s_n\}$ and $\mathbb{T} = \{t_1, \dots, t_n\}$ are $(\theta, c)$-separated from one another, where the values $1-\theta$ and $\delta$ are positive but not small. We can replace $\delta = \delta_{c,\mathbb{S}} = \min_{i=1}^n |s_i - c|$ by $\delta = \delta_{c,\mathbb{T}} = \min_{j=1}^n |t_j - c|$ throughout because of the symmetric roles of the sets $\mathbb{S}$ and $\mathbb{T}$.

The Cauchy matrices with the knot set $\mathbb{T} = \{t_j = b\omega^{j-1},\ j = 1, \dots, n\}$ have link to Vandermonde matrix structure, and we call them *CV matrices*. Their transposes have the knot set $\mathbb{S} = \{s_j = a\omega^{j-1},\ j = 1, \dots, n\}$, have link to transposed Vandermonde matrix, and we call them $CV^T$ matrices.

**Theorem 5.1.** *Assume positive integers $g$, $h$ and $n$, a scalar $e$, and a CV matrix $C = C_{\mathbf{s},e} = (\frac{1}{s_i-t_j})_{i,j=1}^n$ such that $t_j = b(\omega^{j-1})$ for $j = 1, \dots, n$, $gh = n$, $n$ is not small, and $|b| = 1$. Then there is a permutation $n \times n$ matrix $P$ such that $CP$ is a $3 \times g$ block matrix with block columns $(C_{j,-}^T \mid \Sigma_j^T \mid C_{j,+}^T)^T$, $j = 0, \dots, g-1$, where the diagonal blocks $\Sigma_j$ have sizes $n_j \times h$, and the rows of the blocks $\Sigma_j$ and $\Sigma_k$ lie in pairwise distinct sets of rows of the matrix $CP$ unless $|j-k| \le 1$ or $|j-k| = g-1$ (and so the blocks $\Sigma_1, \dots, \Sigma_g$ 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 $n$ knots 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 $\phi_i$ 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 $e = 1$. Write $\mathbb{S}_j = \{s_j\}_j \in \mathbb{S}$ and $\mathbb{T}_j = \{\omega^l\}_{l=jh}^{j(h+1)-1} \in \mathbb{T}$ to 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 $\omega^{jh}$ and $\omega^{(j+1)h}$, 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 $\phi_j$ satisfy $2\pi jh/n \le \phi_j < 2\pi(j(h+1)-1)/n$, $j = 0,\dots,g-1$.

Write $(a \hookrightarrow b)$ to denote the arc of the unit circle $\{z : \ |z| = 1\}$ with the end points $a$ and $b$. For every $j$, $j = 1,\dots,g$, choose a center $c_j$ on the arc $(\omega_{4n}^{(4j+1)h} \hookrightarrow \omega_{4n}^{(4j+3)h})$. This arc has the length $\pi h/n$ and shares the midpoint $\omega_{2n}^{(2j+1)h}$ with the arc $(\omega^{jh} \hookrightarrow \omega^{(j+1)h})$, 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 required). This is possible because the set has exactly $n$ elements. For $j = 0,\dots,g-1$, index by $jh,\dots,j(h+1)-1$ the columns shared by the blocks $C_{j,-}$, $\Sigma_j$ and $C_{j,+}$ and index the rows of the blocks $\Sigma_j$ 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 $\mathbb{T}_k = \{\omega^l\}_{(k-1)h}^{kh-1}$ are $(1/2, c_j)$-separated from one another unless $|j-k| \le 1$ or $|j-k| = g-1$, and this immediately supports the separation requirement of the theorem. □

Apply Corollary 5.1 for $q = 1/2$, $\delta = 0.5h/n^2$, $C = (C_{u,-} \mid C_{u,+})^T$, and $u = 1,\dots,g$ and obtain the following corollary.

**Corollary 5.2.** *The matrix $PC$ of Theorem 5.1 can be represented as*

$$PC = \Sigma + \widehat{C} + E \tag{5.5}$$

*where $\Sigma$ 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* $||E|| \le n^3 2^{2-k}/h$.

The corollary bounds the numerical rank of CV matrices and can be extended to Cauchy matrices of a large class. Indeed Theorem 5.1 and the corollary can be immediately extended to the case where $h$ does not divide $n$ (in this case write $g = \lceil n/h \rceil$) and to the case of $CV^T$ matrices $C = (\frac{1}{s_i - t_j})_{i,j=1}^n$ for $s_i = a\omega^{i-1}$ for all $i$ and $|b| = 1$ because $-C^T$ is the Cauchy matrix $(\frac{1}{t_i - s_j})_{i,j=1}^n$. The proof techniques enable extension to the case where instead of the set $\mathbb{T}$ on the unit circle $\{x : |x| = 1\}$ we have the points of such a set more or less equally spaced either on a line interval, which has a length between 1 and 2 (say) and lies in the complex plane not very far from the origin, or where these points are more or less equally spaced on an approximation of such an interval by a segment of a curve.

# 6 Fast approximate multiplication of CV, $CV^T$, Vandermonde and transposed Vandermonde matrices by a vector

Write $\alpha(M) = \alpha(M,k,h)$ to denote the number of ops sufficient to approximate the product of the matrix $M$ by a vector $\mathbf{v}$ within the error norm bound $n^3 2^{2-k}||\mathbf{v}||/h$. Corollary 5.2 implies that $\alpha(C) \le \alpha(\Sigma) + \alpha(\widehat{C})$.

Defining (5.5) we can choose any positive integer $k$ and any pair of integers $g$ and $h$ such that $1 \le g < n$, $1 \le h < n$ and $gh \ge n$. For a fixed positive $u$ we can ensure the bound $||E|| \le 2^{-u}$ by choosing $k = \lceil 3(u+2)\log_2 n \rceil$, and so for a constant $u$ it is sufficient to choose $k$ of order $\log n$.

In Section 4 we have transformed all Vandermonde matrices into Cauchy matrices of the above class CV, and so we can apply the same cost bounds to approximate Vandermonde multiplication by a vector, except that the bound on the error norm $||E||$ can grow by a factor of $\mu \le 1/\min_i |s_i^n - a^n|$. Having $n$ complex points $s_1^n,\dots,s_n^n$ fixed, we can readily select a point $a^n$ such that $\mu \le 3n$, and then we can still ensure the same bound on the output error norm by increasing $k$ by $\lceil \log_2(3n) \rceil$. This observation extends our results on approximate multiplication by a vector from the class of CV matrices to any Vandermonde matrix $V$, that is to any input of Problem 1. By applying the transposition we can extend these results also to transposed Vandermonde and $CV^T$ matrices.

Clearly $\alpha(\Sigma) \le 6hn - n$ because the matrix $\Sigma$ has at most $3hn$ nonzero entries, whereas $\alpha(\widehat{C}) < 2(n+h)kg$ because $\widehat{C}$ is given by the $g$ products of matrices of sizes at most $n \times h$ having rank

$k$. It follows that we can compute a close approximation to the vectors $C\mathbf{v}$ and $V\mathbf{v}$ within $2^{-b}||\mathbf{v}||$ fast, namely, by using $O(n\sqrt{n\log n})$ ops if we choose $h$ of about $\sqrt{n\log n}$ and choose $g$ of about $\sqrt{n/\log n}$. *This algorithm accelerates the known algorithms for multipoint polynomial evaluation by a factor of* $\sqrt{n/\log n}$, but we will obtain faster algorithms because the upper estimate for $\alpha(\widehat{C})$ above is crude and overly pessimistic.

# 7 HSS matrices and accelerated multiplication by a vector and the solution of a linear system of equations

Hereafer we use the acronym "HSS" for "hierarchically semiseparable". To decrease the bound $\alpha(\widehat{C})$, we will exploit the HSS (that is hierarchically semiseparable) structure of the matrix $\widehat{C}$, whose construction and study are closely related to the Multipole algorithms (cf. [GR87]). The HSS structure was already implicit in our study in the two previous sections.

Next we recall one of the equivalent or nearly equivalent definitions and some fundamental properties of this class, referring the reader to [VVGM05], [MRT05], [VVM07], [VVM08], [X12], and the bibliography therein on the long history of the study of this and similar matrix classes, known under the names of matrices with low Hankel rank, rank structured matrices, quasiseparable, and weakly, recursively, or sequentially semiseparable matrices. See [GR87], [LRT79], [PR93], on the related subjects of Multipole and Nested Dissection Algorithms.

**Definition 7.1.** *A matrix is an $(l,u)$-HSS matrix if $l$ is the maximum rank of its subdiagonal blocks and if $u$ is the maximum rank of its superdiagonal blocks, that is blocks lying strictly below or strictly above the diagonal, respectively.*

Any banded matrix $B$ having the pair $(l,u)$ of lower and upper bandwidths is an $(l,u)$-HSS matrix, and so is its inverse if the matrix is nonsingular. It is well known that a banded $n\times n$ matrix having the pair $(l,u)$ of lower and upper bandwidths can be multiplied by a vector by using $O((l+u)n)$ ops, whereas $O((l+u)^2n)$ ops are sufficient to solve a nonsingular linear system of $n$ equations with such banded matrices. Application of Multipole algorithms enables extension of both properties to $(l,u)$-HSS matrices (see [GR87], [MRT05], [CGS07], [XXG12]). Furthermore, like the matrices of the classes $\mathcal{T}$, $\mathcal{H}$, $\mathcal{V}$ and $\mathcal{C}$, the HSS matrices allow compressed representation: one can define generalized generators that readily express the $n^2$ entries of an $(l,u)$-HSS $n\times n$ matrix 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 $O((l+u)^2n)$ ops. See [XXG12] and references therein on the supporting algorithms and their efficient implementation.

The block diagonal matrix $\Sigma$ has at most $3hn$ entries. The matrix $\widehat{C}$ consists of the off-diagonal blocks. By combining Theorem 5.1 and Corollary 5.2 with the Multipole/HSS techniques of [MRT05], [CGS07], [XXG12], deduce that for any real constant $b$ and for $k=\lceil 3(b+2)\log_2 n\rceil$, the matrix $\widehat{C}$ of (5.5) 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. By applying the cited results on multiplication of HSS matrices by a vector and on the solution of HSS nonsingular linear systems of equations, we obtain the following results (cf. [Pa, Theorem 33]).

**Theorem 7.1.** *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 $O(bn\log n)$ ops provided that $M$ is a CV, $CV^T$, Vandermonde or transposed Vandermonde $n\times n$ matrix. (ii) The op 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 op 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.*

# 8 Extension of matrix structures and complexity estimates

To extend our progress we recall and employ the Sylvester displacements $AM-MB$ of matrices with the structures of Cauchy and Vandermonde types. We say that an $n \times n$ matrix $M = C_{\mathbf{s},\mathbf{t}}(G, H)$ has a $d$-structure of the Cauchy type $\{\mathbf{s}, \mathbf{t}\}$ with a generator $G, H$ if $D_{\mathbf{s}}M - MD_{\mathbf{t}} = GH^T$ for a pair of $n \times d$ matrices $G$ and $H$ with columns $\mathbf{g}_j$ and $\mathbf{h}_j$, for $j = 1, \ldots, d$, and rows $\mathbf{u}_i$ and $\mathbf{v}_i$, for $i = 1, \ldots, n$, respectively, or equivalently if

$$M = \sum_{j=1}^{d} D(\mathbf{g}_j) C D(\mathbf{h}_j) = \left(\frac{\mathbf{u}_i^T \mathbf{v}_j}{s_i - t_j}\right)_{i,j=0}^{n-1} \quad \text{for } C = \left(\frac{1}{s_i - t_j}\right)_{i,j=1}^{n} \tag{8.1}$$

(cf. [P01, Example 1.4.1 and 4.6.4]). In this case we write $M = C_{\mathbf{s},\mathbf{t}}(G, H)$. In particular $M = (\frac{1}{s_i - t_j})_{i,j=1}^n$ if $d = 1$ and if $G = H = (1, \ldots, 1)^T$. Formula (8.1) expresses a matrix $M$ via a nonunique pair of generators $\{G, H\}$ of its displacement, called *displacement generators* of the matrix.

Similar extensions of other classes of structured matrices such as Toeplitz, Hankel, and Vandermonde matrices $M$ and their transposes have been defined as well. Namely these basic classes have been extended to the matrices whose displacements $AM-MB$ have small ranks (called *displacement ranks* of the matrices $M$) for appropriate nonunique pairs of operator matrices $A$ and $B$, similarly to the pair of operator matrices $A = D_{\mathbf{s}}$ and $B = D_{\mathbf{t}}$, which support such an extension of the class of Cauchy matrices $C_{\mathbf{s},\mathbf{t}} = (\frac{1}{s_i - t_j})_{i,j=1}^n$ by virtue of (8.1). Each pair of operator matrices defines a class of matrices with the associated structure. For example, the pairs $(A, B) = (Z_e, Z_f)$ and $(A, B) = (Z_e, Z_f^T)$ for a pair of distinct constants $e$ and $f$ define the classes of matrices with Toeplitz and Hankel structures, respectively, whereas the following operator matrices $A$ and $B$ for any scalar $f$ define an extension of the class of the Vandermonde matrices $V = V_{\mathbf{x}} = (x_i^j)_{i,j=1}^n$,

$$D_{\mathbf{x}}V - VZ_f = (x_i^n - f)_{i=1}^n (0 \mid \ \ldots \ \mid 0 \mid 1). \tag{8.2}$$

The transpose of this equation is basic for extending the class of transposed Vandermonde matrices $V^T$,

$$Z_f^T V^T - V^T D_{\mathbf{x}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} ((x_i^n - f)_{i=1}^n)^T. \tag{8.3}$$

We refer the reader to [BP94], [GO94], [P01, Sections 4.1–4.5], [PW03], and the bibliography therein on expressing various structured matrices via their displacements (this includes matrices with the structures of Toeplitz, Hankel, Vandermonde and transposed Vandermonde types) and to [P01, Section 1.5] on performing arithmetic operations with matrices in terms of their displacements, exemplified by the following simple but basic result.

**Theorem 8.1.** *(Cf. [P90] and[P01, Theorem 1.5.4].) Assume five matrices $A$, $B$, $C$, $M$ and $N$ with compatible sizes. Then*

$$A(MN) - (MN)C = (AM - MB)N + M(BN - NC), \tag{8.4}$$

*and so if $AM - MB = G_M H_M^T$ and $BN - NC = G_N H_N^T$, then $A(MN) - (MN)C = G_{MN} H_{MN}^T$ for $G_{MN} = (G_M \mid MG_N)$ and $H_{MN} = (N^T H_M \mid H_N)$.*

By virtue of this simple theorem one can transform a pair of operator matrices $(A, B)$ into any other pair of operator matrices by means of multiplication with appropriate multipliers. Consequentltly we can transform the classes of matrices with the structures of Toeplitz, Hankel, Vandermonde, transposed Vandermonde, and Cauchy types into each other at will. This simple observation has already lead to substantial algorithmic benefits, because it enabled us to exploit distinct features of various matrix structures, in particular the fact that the matrix structure of Cauchy type is invariant in row and column interchange (in contrast to the structures of Toeplitz and Hankel types) (cf. [GKO95], [G98]) and enables expansion of the matrix entries into Loran's series (unlike the structures of the three other types) (cf. [MRT05], [CGS07], [XXG12]).

The transformation of the matrix structures of Vandermonde into Cauchy type was the basis for our current progress in this paper, and these transforms can be obtained by means of multiplication of a Vandermonde matrix $V_{\mathbf{s}}$ by the inverse $V_{\mathbf{t}}^{-1}$ of a Vandermonde matrix, where we can choose $\mathbf{t} = (\omega^{j-1})_{j=1}^{n}$ to obtain $V_{\mathbf{t}} = \sqrt{n}\Omega$ and $V_{\mathbf{t}}^{-1} = \sqrt{n}\Omega^H$ (cf. equation (4.4)).

More general transforms of matrix structures enable extension of our algorithms to wider classes of matrices. In particular wby engaging the Cauchy and Cauchy-like matrices we can extend our algorithms to for rational multipoint evaluation and interpolation, but at this point our progress becomes more limited because it becomes harder to control propagation of approximation errors (cf. [Pa]).

# 9 Conclusions

Presently we have no estimates for the treshold input sizes for which our algorithms running in nearly linear time outperform their variant of Section 6 and the known algorithms running in quadratic time. Implementation work for our algorithms should prompt their refinements toward decreasing these values.

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. One can be also motivated to seek further applications and extensions of the proposed techniques, e.g., to operations with confluent Vandermonde matrices, to the computations with Loewner matrices, and to various problems of rational interpolation such as the Nevanlinna–Pick and matrix Nehari problems, where, however, the progress can be limited to the case of sufficiently well conditioned inputs.

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

# References

[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.

[B-O83] M. Ben-Or, Lower Bounds for Algebraic Computation Trees, *Proceedings of 15th Annual ACM Symposium on Theory of Computing (STOC'83)*, 80–86, ACM Press, New York, 1983.

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

[BP70] A. Björck, V. Pereyra, Solution of Vandermonde Systems of Equations, *Math. of Computation*, **24**, 893–903, 1970.

[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.

[F72] C. M. Fiduccia, Polynomial Evaluation via the Division Algorithm: The Fast Fourier Transform Revisited, *Proc. 4th Annual ACM Symp. on Theory of Computing (STOC'72)*, 88–93, 1972.

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

[GG03] J. von zur Gathen, J. Gerhard, *Modern Computer Algebra*, Cambridge University Press, Cambridge, UK, 2003 (second edition).

[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.

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

[H72] E. Horowitz, A Fast Method for Interpolation Using Preconditioning, *Information Processing Letters*, **1, 4**, 157–163, 1972.

[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.

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

[MB72] R. Moenck, A. Borodin, Fast Modular Transform via Division, *Proceedings of 13th Annual Symposium on Switching and Automata Theory*, 90–96, IEEE Computer Society Press, Washington, DC, 1972.

[MRT05] P. G. Martinsson, V. Rokhlin, M. Tygert, A fast algorithm for the inversion of general Toeplitz matrices, em Comput. Math. Appl., **50**, 741–752, 2005.

[P64] F. Parker, Inverses of Vandermonde matrices, *Amer. Math. Monthly*, **71**, 410–411, 1964.

[P66] V. Y. Pan, On Methods of Computing the Values of Polynomials, *Uspekhi Matematicheskikh Nauk*, **21**, **1(127)**, 103–134, 1966. [Transl. *Russian Mathematical Surveys*, **21, 1(127)**, 105–137, 1966.]

[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.

[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, Transformations of Matrix Structures Work Again, Tech. Report TR 2013004, *PhD Program in Comp. Sci., Graduate Center, CUNY*, 2013

Available at http://www.cs.gc.cuny.edu/tr/techreport.php?id=446

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

[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.

[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.

[S73] V. Strassen, Die Berechnungskomplexetät von elementarysymmetrischen Funktionen und von Interpolationskoeffizienten, *Numerische Mathematik*, **20, 3**, 238–251, 1973.

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

[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.

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