Fast Estimates of Hankel Matrix Condition Numbers and Numeric Sparse Interpolation*

Erich L. Kaltofen
Department of Mathematics
North Carolina State University
Raleigh, NC 27695-8205, USA
kaltofen@math.ncsu.edu www.kaltofen.us

Wen-shin Lee
Mathematics and Computer Science Department
University of Antwerp
2020 Antwerpen, Belgium
wen-shin.lee@ua.ac.be

Zhengfeng Yang
Shanghai Key Laboratory of Trustworthy Computing
East China Normal University
Shanghai 200062, China
zfyang@sei.ecnu.edu.cn

ABSTRACT

We investigate our early termination criterion for sparse polynomial interpolation when substantial noise is present in the values of the polynomial. Our criterion in the exact case uses Monte Carlo randomization which introduces a second source of error. We harness the Gohberg-Semencul formula for the inverse of a Hankel matrix to compute estimates for the structured condition numbers of all arising Hankel matrices in quadratic arithmetic time overall, and explain how false ill-conditionedness can arise from our randomizations. Finally, we demonstrate by experiments that our condition number estimates lead to a viable termination criterion for polynomials with about 20 non-zero terms and of degree about 100, even in the presence of noise of relative magnitude 10^{-5}.

Categories and Subject Descriptors

F.2.1 [**Numerical Algorithms and Problems**]: Computations on matrices, Computations on polynomials; G.1.1 [**Interpolation**]; I.1.2 [**Algorithms**]: Algebraic algorithms, Analysis of algorithms

General Terms

Algorithm, Theory

*This material is based on work supported in part by the National Science Foundation under Grant CCF-0830347 (Kaltofen) and a 2010 Ky and Yu-Fen Fan Fund (Kaltofen and Yang).
This material is supported by the Chinese National Natural Science Foundation under Grant 10901055 and the Shanghai Natural Science Foundation under Grant 09ZR1408800 (Yang).

Keywords

Gohberg-Semencul formula, Hankel system, Vandermonde system, sparse polynomial interpolation, early termination

1. INTRODUCTION

The fast solution of Toeplitz- and Hankel-like linear systems is based on the classical 1947 Levinson-Durbin problem of computing linear generators of the sequence of entries. When with exact arithmetic on the entries a singular leading principal submatrix is encountered, look-ahead can be performed: this is the genesis of the 1968 Berlekamp/Massey algorithm. In the numeric setting, the notion of ill-conditionedness substitutes for exact singularity, and the algorithms need to estimate the condition numbers of the arising submatrices. By a result of Siegfried Rump [22], the structured condition number of an $n \times n$ Toeplitz/Hankel matrix H is equal to its unstructured condition number $\kappa_2(H)$: we have a Hankel perturbation matrix ΔH with spectral (matrix-2) norm $\|\Delta H\|_2 = 1/\|H^{-1}\|_2 = \|H\|_2/\kappa_2(H)$ such that $H + \Delta H$ is singular [22, Theorem 12.1]. Thus Trench's [23] (exact) inverse algorithm can produce upper and lower bounds for those distances from the Frobenius (Euclidean-vector-2) norm estimates for $1/\sqrt{n}\,\|H^{-1}\|_F \le \|H^{-1}\|_2 \le \|H^{-1}\|_F$. The look-ahead algorithms, such as the Berlekamp/Massey solver, however, need such estimates for all $k \times k$ leading principal submatrices $H^{[k]}$ for $k = 1, 2, \ldots$ of an unbounded Hankel matrix. Here we show how the 1972 Gohberg-Semencul "off-diagonal" (JL)U-representations [17, 7] can give estimates for all $\kappa_2(H^{[k]})$ in quadratic arithmetic overall time. Although our estimates are quite accurate (see Figures 1 and 2), the computation of the actual condition numbers of all leading principal submatrices in quadratic arithmetic time remains an intriguing open problem. We note that fraction-free implementations [2, 21] of the Berlekamp/Massey algorithms and inverse generators only produce values of polynomials in the entries, and our numeric Schwartz/Zippel lemma [20, Lemma 3.1] is applicable. In Section 3 we will discuss the numeric sensitivity of the expected values in that Lemma.

Our investigations are motivated by numeric sparse polynomial interpolation algorithms [11, 12, 20, 13]. In the Ben-Or/Tiwari version of this algorithm, one first computes for a sparse polynomial $f(x) \in \mathbb{C}[x]$ the linear generator for $h_l = f(\omega^{l+1})$ for $l = 0, 1, 2, \ldots$ In some exact as well as numeric algorithms, ω is selected as a random p-th root of unity. Then by the theory of early termination of that algorithm [18], the first $t \times t$ leading principle submatrix $H^{[t]}$ in the (infinite) Hankel matrix H with entries in row i and column j, $(H)_{i,j} = h_{i+j-2}$ where $k = 1, 2, \ldots, t$,

$$H^{[k]} = \begin{bmatrix} h_0 & h_1 & h_2 & h_3 & \ldots & h_{k-1} \\ h_1 & h_2 & h_3 & h_4 & \iddots & h_k \\ h_2 & h_3 & h_4 & h_5 & \iddots & h_{k+1} \\ h_3 & h_4 & h_5 & h_6 & \iddots & \vdots \\ \vdots & \iddots & \iddots & \iddots & \iddots & \\ h_{k-1} & h_k & h_{k+1} & \ldots & & h_{2k-2} \end{bmatrix} \quad (1)$$

that is singular has with high probability dimension t = the number of non-zero terms in f. For technical reasons in proof of the probabilistic analysis, we must skip over the value $f(\omega^0) = f(1)$. Hence, by the previously described algorithm for lower and upper bounds of the condition numbers of $H^{[k]}$, we can determine the number of terms t.

The Berlekamp/Massey algorithm solves the problem of computing Padé forms along a diagonal of the Padé table. Cabay and Meleshko gave an algorithm to bound from above the condition numbers of all the principal submatrices of the associated Hankel system in the well-conditioned case [8], which has been further refined and extended, for example, see [1, 6]. Here we focus less on stably computing the Gohberg-Semencul updates (see those cited papers—we require no look-ahead), but on fast and accurate lower and upper bounds of the condition numbers, which constitute the early termination criterion. We add that our approach here, namely testing submatrices for ill-conditionedness, is problematic for the approximate GCD problem (cf. [5, 4]): we know now that Rump's property is false for Sylvester matrices: the unstructured condition number of a Sylvester matrix can be large while the structured condition number is small (see [19, Example 4.2]).

We approach the problem of incrementally estimating Hankel submatrices through a numerical interpretation of the Berlekamp/Massey algorithm [21] that combines with an application of the numerical Schwartz/Zippel lemma [20]. We report some experimental results of our implementation in Section 4. For a 20 term polynomial of degree 100 the algorithm correctly computes $t = 20$ in quadratic time (given the polynomial evaluations h_l) (see Table 1). There are several issues to scrutinize. First, we only have bounds on the condition numbers and must verify that our estimates are sufficiently accurate. One would expect since for $h_l = f(\omega^{l+1})$ the Hankel matrix $H^{[k]}$ in (1) has additional structure that the upper bound estimates for the Rump condition numbers are good indicators of numeric non-singularity. Second, early termination is achieved by randomization, whose probabilistic analysis, namely the separation of the decision quantities, i.e., determinants or condition numbers, from very small values (in terms of absolute values), applies to exact computation. We can relate the numeric sensitivity of those decision quantities via a factorization of $H^{[t]}$ to clustering of term values on the unite circle, which is partially overcome by evaluation at several random ω simultaneously.

Throughout the paper, we will use the boldfaced letters $\mathbf{x}$ and $\mathbf{y}$ to denote the vectors $[x_1, \ldots, x_n]^T$, $[y_1, \ldots, y_n]^T$, respectively; $H^{[k]}$ denotes the $k \times k$ leading principle submatrix; $\kappa(H) = \|H\| \cdot \|H^{-1}\|$ denotes the unstructured condition number of H.

2. CONDITION NUMBER ESTIMATE

Suppose a $n \times n$ Hankel matrix H is strongly regular, i.e., all the leading principle submatrices of H are nonsingular. Following an efficient algorithm for solving Hankel systems [14], we present a recursive method to estimate the condition numbers of all leading principal Hankel submatrices in quadratic time.

For simplicity, in this section we only consider the condition number of the Hankel matrix with respect to 1 norm, i.e., $\kappa_1(H) = \|H\|_1 \cdot \|H^{-1}\|_1$, since all other operator norms can be bounded by the corresponding 1 norm.

In order to estimate the condition numbers of all leading principal submatrices, we need to estimate the norms of all leading principal submatrices as well as their inverses. Since a Hankel matrix H has recursive structure, the norm of $H^{[i+1]}$ can be obtained in $O(i)$ flops if the norm of $H^{[i]}$ is given, which implies that computing the norms of all leading principal submatrices can be done in $O(n^2)$ flops. Therefore, our task is to demonstrate how to estimate the norm of $(H^{[i+1]})^{-1}$ in linear time if one has the estimate of the norm of $(H^{[i]})^{-1}$.

As restated in Theorem 1, the Gohberg-Semencul formula for inverting a Hankel matrix plays an important role in our method. Note that for a given Hankel matrix H, JH is a Toeplitz matrix, where J is an anti-diagonal matrix with 1 as its nonzero entries.

Theorem 1 *Given a Hankel matrix H, suppose $\mathbf{x}$ and $\mathbf{y}$ are the solution vectors of the systems of the equations $H\mathbf{x} = e_1$, $H\mathbf{y} = e_n$, where $e_1 = [1, 0, 0, \ldots, 0]^T$ and $e_n = [0, \ldots, 0, 1]^T$. Then if $x_n \neq 0$ the inverse H^{-1} satisfies the identity*

$$H^{-1} = \frac{1}{x_n} \underbrace{\begin{bmatrix} x_1 & x_2 & \cdot & x_{n-1} & x_n \\ x_2 & x_3 & \cdot & x_n & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{n-1} & x_n & \cdot & 0 & 0 \\ x_n & 0 & \cdot & 0 & 0 \end{bmatrix}}_{JL_1} \underbrace{\begin{bmatrix} y_1 & y_2 & \cdot & y_{n-1} & y_n \\ 0 & y_1 & \cdot & y_{n-2} & y_{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & y_1 & y_2 \\ 0 & 0 & \cdot & 0 & y_1 \end{bmatrix}}_{R_1} - \frac{1}{x_n} \underbrace{\begin{bmatrix} y_2 & y_3 & \cdot & y_n & 0 \\ y_3 & y_4 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ y_n & 0 & \cdot & 0 & 0 \\ 0 & 0 & \cdot & 0 & 0 \end{bmatrix}}_{JL_2} \underbrace{\begin{bmatrix} 0 & x_1 & \cdot & x_{n-2} & x_{n-1} \\ 0 & 0 & \cdot & x_{n-3} & x_{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 0 & x_1 \\ 0 & 0 & \cdot & 0 & 0 \end{bmatrix}}_{R_2} \quad (2)$$

The inversion formula of a Hankel matrix leads to a polynomial form, whose coefficients involve vectors $\mathbf{x}$ and $\mathbf{y}$.

Theorem 2 *Under the assumptions of Theorem 1, suppose $H^{-1} = [\gamma_1, \gamma_2, \ldots, \gamma_n]$, where γ_k denotes the k-th column of H^{-1}. Then we have*

$$\gamma_k = [\mathbf{v}_{n-1}, \mathbf{v}_{n-2}, \ldots, \mathbf{v}_0]^T,$$

where $\mathbf{v}_i$ denotes the coefficient of the polynomial $(1/x_n) \cdot$

$(\phi \cdot g_k - u_k \cdot v)$ *with respect to* Z^i, *where*

$$\phi = \sum_{i=1}^{n} x_i Z^{n-i}, \quad g_k = \sum_{i=1}^{k} y_i Z^{k-i},$$

$$v = \sum_{i=2}^{n} y_i Z^{n-i}, \quad u_1 = 0, \quad u_k = \sum_{i=1}^{k-1} x_i Z^{k-i}, k = 2, \ldots, n.$$

PROOF. Considering the polynomial product $\phi \cdot g_k$, we obtain the coefficient subvector

$$\mu = [\psi_{n-1}, \psi_{n-2}, \ldots, \psi_0]^T,$$

where ψ_j is the coefficient of $\phi \cdot g_k$ with respect to Z^j. Expanding $\phi \cdot g_k$, we get

$$\psi_j = \begin{cases} \sum_{i=0}^{j} x_{n-i} y_{k+i-j}, & \text{if } 0 \le j \le k, \\ \sum_{i=0}^{k-1} x_{n+i-j} y_{k-i}, & \text{if } k < j < n. \end{cases}$$

Observing the formula of H^{-1} in Theorem 1, We find that the coefficient subvector

$$\mu = JL_1.[y_k, y_{k-1}, \ldots, 0, 0]^T, \tag{3}$$

where $[y_k, y_{k-1}, \ldots, 0, 0]$ is the k-th column of R_1. Since $u_1 = 0$, it is obvious that the coefficient vector of $u_1 \cdot v$ is a zero vector. Considering the polynomial product $u_k \cdot v, 2 \le k \le n$, we obtain the coefficient subvector

$$\nu = [\chi_{n-1}, \chi_{n-2}, \ldots, \chi_0]^T,$$

where χ_j is the coefficient of $u_k \cdot v$ with respect to Z^j. Expanding $u_k \cdot v$, we get

$$\chi_j = \begin{cases} 0 & \text{if } j = 0 \\ \sum_{i=0}^{j} y_{n-i} x_{k-j+i}, & \text{if } j \le k-1, \\ \sum_{i=1}^{k-1} y_{n+1-j} x_{k-i}, & \text{if } k \le j < n. \end{cases}$$

The coefficient subvector

$$\nu = JL_2.[x_{k-1}, x_{k-2}, \ldots, 0, 0]^T, \tag{4}$$

where $[x_{k-1}, x_{k-2}, \ldots, 0, 0]^T$ is the k-th column of R_2. According to (3) and (4), it can be verified that the column γ_k is able to be represented by the coefficients of $(1/x_n) \cdot (\phi \cdot g_k - u_k \cdot v)$. □

Corollary 1 *Given a nonsingular Hankel matrix* H, *suppose* $\mathbf{x}$ *and* $\mathbf{y}$ *are the solutions of the equations of* $H\mathbf{x} = e_1$ *and* $H\mathbf{y} = e_n$, *then*

$$\|\mathbf{x}\|_1 \le \|H^{-1}\|_1 \le \frac{2\|\mathbf{x}\|_1 \cdot \|\mathbf{y}\|_1}{|x_n|}. \tag{5}$$

PROOF. Since $x_n = y_1$, (5) can be concluded from Theorem 2. □

Given a strongly regular matrix H, we discuss how to use the bound in (5) to estimate $\|(H^{[i]})^{-1}\|_1$ for all the leading principal submatrices in quadratic time. A linear time recursive algorithm is presented in [14] to obtain the solution vector of $H^{[i+1]}\mathbf{x}_{i+1} = e_{i+1}$ from the solution vector of $H^{[i]}\mathbf{y}_i = e_i$, where $H^{[i]}$ and $H^{[i+1]}$ are the $i \times i$ and $(i+1) \times (i+1)$ leading principal submatrices of H respectively, and $e_i = [0, \ldots, 0, 1]^T$, $e_{i+1} = \left[0, e_i\right]^T$. Moreover, according to the Lemma in [14], for an $i \times i$ strongly regular Hankel matrix H, the solution $\mathbf{x}$ of the linear system $H\mathbf{x} = e_1$ can be obtained in $O(i)$ flops if one has the solution $H\mathbf{y} = e_i$. In other words, for a given strongly regular $k \times k$ Hankel matrix H, all of the solution vectors $\mathbf{x}$ and $\mathbf{y}$ of $H^{[i]}\mathbf{x} = e_1, H^{[i]}\mathbf{y} = e_i, i = 1, 2, \ldots, k$ can be obtained in $O(k^2)$ flops. Hence the recursive algorithm presented in [14] is applicable to obtain the bounds (5) of all $\|(H^{[i]})^{-1}\|_1$ in quadratic time.

Theorem 3 *Under the same assumptions as above, given a strongly regular* $n \times n$ *Hankel matrix* H, *we have*

$$\|\mathbf{x}^{[i]}\|_1 \|H^{[i]}\|_1 \le \kappa_1(H^{[i]}) \le \frac{2\|\mathbf{x}^{[i]}\|_1 \|\mathbf{y}^{[i]}\|_1}{|x_i^{[i]}|} \|H^{[i]}\|_1, \tag{6}$$

where $\mathbf{x}^{[i]}$ *and* $\mathbf{y}^{[i]}$ *are the solution vectors of* $H^{[i]}\mathbf{x}^{[i]} = e_1$ *and* $H^{[i]}\mathbf{y}^{[i]} = e_i$, *and* $x_i^{[i]}$ *denotes the last element of* $\mathbf{x}^{[i]}$. *Furthermore, computing the lower bounds and the upper bounds* (6) *of all* $\kappa_1(H^{[i]})$, $1 \le i \le n$, *can be achieved in* $O(n^2)$ *arithmetic operations.*

PROOF. We conclude (6) from Corollary 1. As discussed, $\|H^{[i]}\|_1$ and the bounds of $\|(H^{[i]})^{-1}\|_1$ can be obtained in quadratic time, which means that the computation of the condition number bounds (6) is $O(n^2)$. □

3. EARLY TERMINATION IN NUMERICAL SPARSE INTERPOLATION

We apply our method to the problem of early termination in numerical sparse polynomial interpolation. Consider a black box univariate polynomial $f \in \mathbb{C}[x]$ represented as

$$f(x) = \sum_{j=1}^{t} c_j x^{d_j}, \quad 0 \ne c_j \in \mathbb{C} \tag{7}$$

and $d_j \in \mathbb{Z}_{\ge 0}, d_1 < d_2 < \cdots < d_t$.

Suppose the evaluations of $f(x)$ contain added noise. Let δ be a degree upper bound of f, which means, $\delta \ge \deg(f) = d_t$. Our aim is to determine the number of terms t in the target polynomial f.

In exact arithmetic, the early termination strategy [18] can determine the number of terms with high probability. Choose a random element ω from a sufficiently large finite set and evaluate $f(x)$ at the powers of ω

$$h_0 = f(\omega), h_1 = f(\omega^2), h_2 = f(\omega^3), \ldots.$$

Consider a $(\delta+1) \times (\delta+1)$ Hankel matrix $H = (H)_{i,j} = h_{i+j-2}$ as in (1). The early termination algorithm makes, with high probability, all principal minors $H^{[j]}$ non-singular for $1 \le j \le t$ [18]. Moreover, $H^{[j]}$ must be singular for all $t < j \le \delta$. Hence the number of terms t is detected as follows: for $j = 1, 2, \ldots$, the first time $H^{[j]}$ becomes singular is when $j = t + 1$.

In the numerical setting, such a singular matrix is often extremely ill-conditioned. So we need to measure the singularity for a given numerical Hankel matrix. In [22], it is proved that for a Hankel matrix the structured condition number with respect to the spectral norm is equivalent to the regular condition number. In other words, for a given Hankel matrix $H^{[j]}$, the distance measured by spectral norm to the nearest singular Hankel matrix is equivalent to $1/\|(H^{[j]})^{-1}\|_2$. In the following we investigate how to estimate the spectral norm of the inverse of a Hankel matrix.

Theorem 4 *Using the same quantities as above, we have*

$$\frac{\|\mathbf{x}\|_1}{\sqrt{n}} \leq \|H^{-1}\|_2 \leq \frac{2\|\mathbf{x}\|_1 \cdot \|\mathbf{y}\|_1}{|x_n|}. \quad (8)$$

PROOF. Given a $n \times n$ arbitrary matrix A, we have the following bound of $\|A\|_2$ [16]

$$\frac{\max\{\|A\|_1, \|A\|_\infty\}}{\sqrt{n}} \leq \|A\|_2 \leq \sqrt{\|A\|_1\|A\|_\infty}.$$

According to Corollary 1 and $\|H^{-1}\|_\infty = \|H^{-1}\|_1$, it can be verified that $\|H^{-1}\|_2$ satisfies (8). □

We following the bound of $\|H^{-1}\|_2$ in (8) and present a recursive algorithm to detect the sparsity of the interpolating polynomial. Suppose we are given an approximate black box polynomial of f represented as (7), its degree bound δ and a chosen tolerance τ, our aim is to recover the number of terms t. We proceed our recursive algorithm as the following. First choose a random root of unity ω with a prime order $p \geq \delta$ and obtain the approximate evaluations

$$h_\ell \approx f(\omega^{\ell+1}), \quad \ell = 0, 1, 2, \ldots.$$

Then we compute iteratively for $k = 1, 2, \ldots$ and $H^{[k]} = [h_{i+j-2}]$ the vectors $\mathbf{x}^{[k]}$ and $\mathbf{y}^{[k]}$ as stated in [14], where $\mathbf{x}^{[k]}$ and $\mathbf{y}^{[k]}$ are the solutions of $H^{[k]}\mathbf{x}^{[k]} = e_1$ and $H^{[k]}\mathbf{y}^{[k]} = e_k$. We break out of the loop until

$$\frac{|x_k^{[k]}|}{2\|\mathbf{x}^{[k]}\|_1 \cdot \|\mathbf{y}^{[k]}\|_1} \leq \tau,$$

which means that for all $j = 1, 2, \ldots, k-1$ the distance between $H^{[j]}$ and the nearest singular Hankel matrix is greater than the given tolerance τ. This is because

$$\frac{1}{\|(H^{[k]})^{-1}\|_2} \geq \frac{|x_k^{[k]}|}{2\|\mathbf{x}^{[k]}\|_1 \cdot \|\mathbf{y}^{[k]}\|_1} > \tau.$$

At this stage, we claim that $H^{[k-1]}$ is strongly regular and obtain $t = k - 1$ as the number of the terms in f.

In sparse polynomial interpolation, the associated Hankel matrix $H^{[t]}$ has additional structure that allows a factorization

$$H^{[j]} = \begin{bmatrix} h_0 & h_1 & \ldots & h_{j-1} \\ h_1 & h_2 & \ldots & h_j \\ \vdots & \vdots & \ddots & \vdots \\ h_{j-1} & h_j & \ldots & h_{2j-2} \end{bmatrix}$$
$$= \underbrace{\begin{bmatrix} 1 & 1 & \ldots & 1 \\ b_1 & b_2 & \ldots & b_t \\ \vdots & \vdots & \ddots & \vdots \\ b_1^{j-1} & b_2^{j-1} & \ldots & b_t^{j-1} \end{bmatrix}}_{V^{[j]}} \underbrace{\begin{bmatrix} c_1b_1 & 0 & \ldots & 0 \\ 0 & c_2b_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & c_tb_t \end{bmatrix}}_{D}$$
$$\times \underbrace{\begin{bmatrix} 1 & b_1 & \ldots & b_1^{j-1} \\ 1 & b_2 & \ldots & b_2^{j-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & b_t & \ldots & b_t^{j-1} \end{bmatrix}}_{(V^{[j]})^{\mathrm{T}}}, \quad (9)$$

in which $b_k = \omega^{d_k}$ for $1 \leq k \leq t$.

When $j = t$, the condition of the associated Hankel system is linked to the embedded Vandermonde system [11, 12]

$$\|(V^{[t]})^{-1}\|^2 \cdot \max_j \frac{1}{|c_j|} \geq \|(H^{[t]})^{-1}\| \geq \frac{\|(V^{[t]})^{-1}\|^2}{\sum_{1\leq j\leq t}|c_j|} \quad (10)$$

(see Proposition 4.1 in [12].)

Since all b_j are on the unit circle, according to [9] the inverse of the Vandermonde matrix $\|V^{-1}\|$ can be bounded by

$$\begin{aligned} \|(V^{[t]})^{-1}\| &\leq \max_{1\leq j\leq t} \prod_{j=1, j\neq k}^{t} \frac{1+|b_j|}{|b_j - b_k|} \\ &= \max_{1\leq j\leq t} \frac{2^{t-1}}{\prod_{j\neq k}|b_j - b_k|}. \end{aligned} \quad (11)$$

As for the lower bound, $\|(V^{[t]})^{-1}\|_\infty = 1$ achieves the optimal condition if b_j are evenly distributed on the unit circle (see Example 6.4 in [10].)

In [11, 12], the sparsity t is given as input. So the randomization is used to improve the condition of the overall sparse interpolation problem, where the recovery of both the non-zero terms and the corresponding coefficients depend on the condition of the embedded Vandermonde system. While [11, 12] further exploit a generalized eigenvalue reformulation [15] in sparse interpolation, interestingly the condition of the associated generalized eigenvalue problem is still dependent on the condition of the embedded Vandermonde system [12, Theorem 4.2] (see [3] for further analysis and discussion.)

Now our purpose now to detect the number of terms. In other words, we intend to use the estimated conditions to determine the number of terms t. (After t is determined, for polynomial interpolation one still needs to recover the t exponents $d_1, \ldots d_t$ and the associated coefficients $c_1, \ldots, c_t$ in f.)

Recall that in exact arithmetic, $H^{[j]}$ is singular for $j > t$. In a numerical setting, such $H^{[j]}$ is expected to be extremely ill-conditioned. If $H^{[t]}$ is relatively well-conditioned, then one can expect a surge in the condition number for $H^{[t+1]}$. We apply the randomization idea in [11, 12] to obtain the heuristic of achieving relatively well-conditioned $H^{[t]}$ with high probability.

Following (11), the distribution of $b_1, \ldots, b_t$ on the unit circle affects the upper bound on the condition of the Vandermonde system $V^{[t]}$. Let $b_j = e^{2\pi i d_j/p}, b_k = e^{2\pi i d_k/p}$ for a prime $p \geq \delta$. Both b_j and b_k are on the unit circle. Let $\Delta_{jk} = |d_j - d_k| \geq 1$, the distance between b_j and b_k is

$$\begin{aligned} |b_j - b_k| &= \sqrt{(1-\cos(2\pi\Delta_{jk}/p))^2 + \sin^2(2\pi\Delta_{jk}/p)} \\ &= \sqrt{2 - 2\cos(2\pi\Delta_{jk}/p)}. \end{aligned}$$

For a fixed t, an upper bound of $\|(V^{[t]})^{-1}\|^2$ is determined by

$$\min_{1\leq j\leq t} \prod_{j\neq k} \left(2 - 2\cos\left(\frac{2\pi\Delta_{jk}}{p}\right)\right), \quad (12)$$

in which p is a chosen prime order for the primitive roots of unity.

According to (12), the corresponding Hankel system $H^{[t]} = V^{[t]}D(V^{[t]})^{\mathrm{T}}$ can be better conditioned if

1. p is smaller; and/or

2. the most (or all) of Δ_{jk} do not belong to smaller values.

In order to achieve a better condition, choosing a larger p would require larger Δ_{jk} and may require a higher precision. Thus we prefer a smaller p. Moreover, in general a smaller p does not affect too much on the overall random distribution of b_j on the unit circle (see Example 3 in Section 4.)

Remark: In the exact case, it may still happen that $H^{[j]}$ is singular for $j \le t$, which corresponds to an ill-conditioned $H^{[j]}$ for $j \le t$ in a finite precision environment. Therefore we perform our algorithm ζ times in the numerical setting. Given $\zeta \in \mathbb{Z}_{>0}$, we choose different random roots of unity $\omega_1, \omega_2, \ldots, \omega_\zeta$. For each j we perform our algorithm on $f(\omega_j^k)$ for $k = 1, 2, \ldots$ and obtain the number of terms as t_j. At the end, we determine $t = \max\{t_1, t_2, \ldots, t_\zeta\}$ as the number of the terms of f.

4. EXPERIMENTS

Our numerical early termination is tested for determining the number of the terms. We set *Digits* := 15 in Maple 13. Our test polynomials are constructed with random integer coefficients in the range between -10 and 10, and with a random degree and random terms in the given ranges.

For each given noise range, we test 50 random black box polynomials and report the times of failing to correctly estimate the number of the terms. In Table 1, *Deg. Range* and *Term Range* record the range of the degree and the number of the term in the polynomials; *Random Noise* records range of random noise added to the evaluations of the black box polynomial; *Fail* reports the times of failing to estimate the correct number of the terms in the target polynomials.

Ex.	*Random Noise*	*Term Range*	*Deg. Range*	*Fail*
1	$10^{-6} \sim 10^{-5}$	$10 \sim 15$	$100 \sim 150$	3
2	$10^{-7} \sim 10^{-6}$	$15 \sim 20$	$100 \sim 150$	1
3	$10^{-8} \sim 10^{-7}$	$20 \sim 25$	$100 \sim 150$	1
4	$10^{-9} \sim 10^{-8}$	$20 \sim 25$	$100 \sim 150$	1

Table 1: Numerical early termination: number of failures out of 50 random polynomials.

Example 1 Given a polynomial f with $\deg(f) = 100$ and the number of term 20, denoted by $T(f) = 20$. Its coefficients are randomly generated as integers between -10 and 10. We construct the black box that evaluates f with added random noises in the range $10^{-9} \sim 10^{-8}$. From such black box evaluations, we generated 1000 random Hankel matrices and compare the condition number $\kappa_1(H) = \|H\|_1 \cdot \|H^{-1}\|_1$ with their corresponding lower bounds $\kappa_{\text{low}}(H)$ and the upper bound $\kappa_{\text{up}}(H)$, shown in (6).

Let $\omega_j = \exp(2 s_j \pi \boldsymbol{i}/p_j) \in \mathbb{C}$ be random roots of unity, where $\boldsymbol{i} = \sqrt{-1}$, p_j prime numbers in the range $100 \le p_j \le 1000$, and s_j random integers with $1 \le s_j < p_j$. We compute the evaluations of the black box for different root of unity ω_j, and construct the associated Hankel matrix $H^{[j]}$. Since $T(f) = 20$, we do 41 evaluations for each j to construct the Hankel matrix with $Dim(H^{[j]}) = 21$:

$$h_{j,l} \approx f(\omega_j^l) \in \mathbb{C}, \quad l = 1, 2, 3, \ldots, 41.$$

We use $H_j^{[k]}$ to denote the $k \times k$ leading principle submatrix of H_j and let

$$r_{\text{low}} = \frac{\|H\|_1 \cdot \|H^{-1}\|_1}{\kappa_{\text{low}}(H)}, \quad r_{\text{up}} = \frac{\kappa_{\text{up}}(H)}{\|H\|_1 \cdot \|H^{-1}\|_1}.$$

For $j = 1, 2, \ldots, 1000$, we obtain the ratios r_{low} and r_{up} for all $k \times k$ leading principle submatrices $H_j^{[k]}$, where $k = 1, 2, \ldots, 20$. Therefore, r_{low} and r_{up} are used to measure whether the lower and upper bounds are close to the actual condition number. We use the histogram to measure the ratios r_{low} and r_{up}. For instance, in Figure 1 (a) and (b) show the distribution of the ratios r_{low} and r_{up} for all 5×5 leading principle submatrices $H_j^{[5]}, j = 1, 2, \ldots, 1000$. In Figure 2, (a) and (b) show the distribution of the ratios r_{low} and r_{up} of all 20×20 leading principle submatrices $H_j^{[20]}, j = 1, 2, \ldots, 1000$.

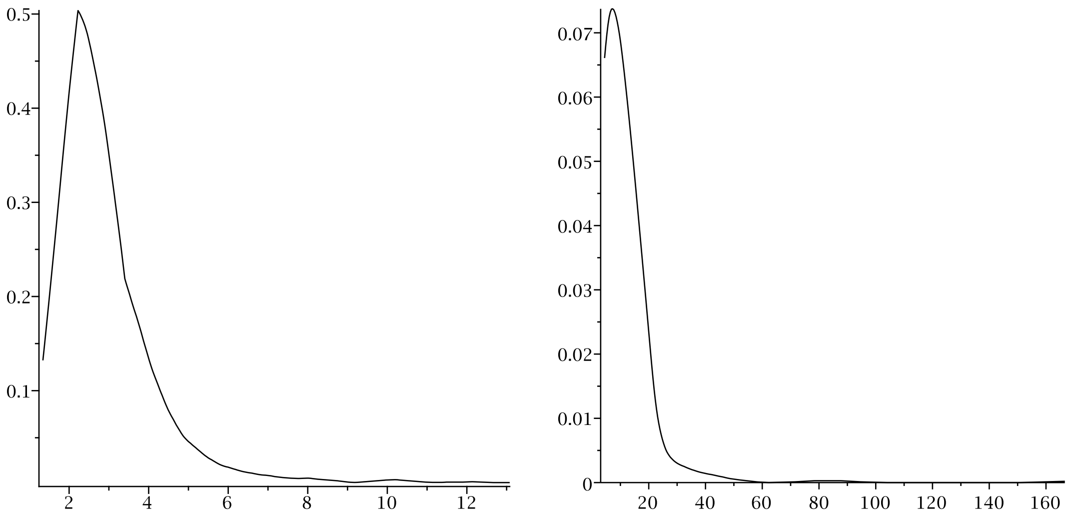
Example 2 We performed $m = 100$ random $n \times n$ Hankel matrices for $n = 4, 8, \ldots, 1024$ whose entries are randomly chosen and follow uniformly probabilistic distribution over $\{c \mid -1 \le c \le 1\}$. Table 2 displays the average value of r_{low} and r_{up}. Here n is the dimension of random Hankel matrices; r_{low} and r_{up} are the average ratios of $\frac{\|H\|_1 \cdot \|H^{-1}\|_1}{\kappa_{\text{low}}(H)}$ and $\frac{\kappa_{\text{up}}(H)}{\|H\|_1 \cdot \|H^{-1}\|_1}$; $\kappa_1 = \|H\|_1 \cdot \|H^{-1}\|_1$ is the average value of the actual condition number obtained for *Digits* := 15 in Maple 13. From the columns of r_{low} and r_{up}, we observe that the upper bound of the condition number depends on the dimension of the Hankel matrices. But the lower bound of the condition number does not seem to be affected much by the increasing of the dimension.

Ex.	n	r_{low}	r_{up}	$\kappa_1(H)$
1	4	3.1431	17.402	523.10
2	8	4.7675	14.023	69.965
3	16	6.4865	30.349	136.59
4	32	11.154	106.35	734.55
5	64	15.691	107.19	1996.1
6	128	24.120	220.42	13218
7	256	34.496	549.59	16932
8	512	49.425	1516.1	16783
9	1024	71.394	1385.0	111375

Table 2: Algorithm performance on condition number estimate

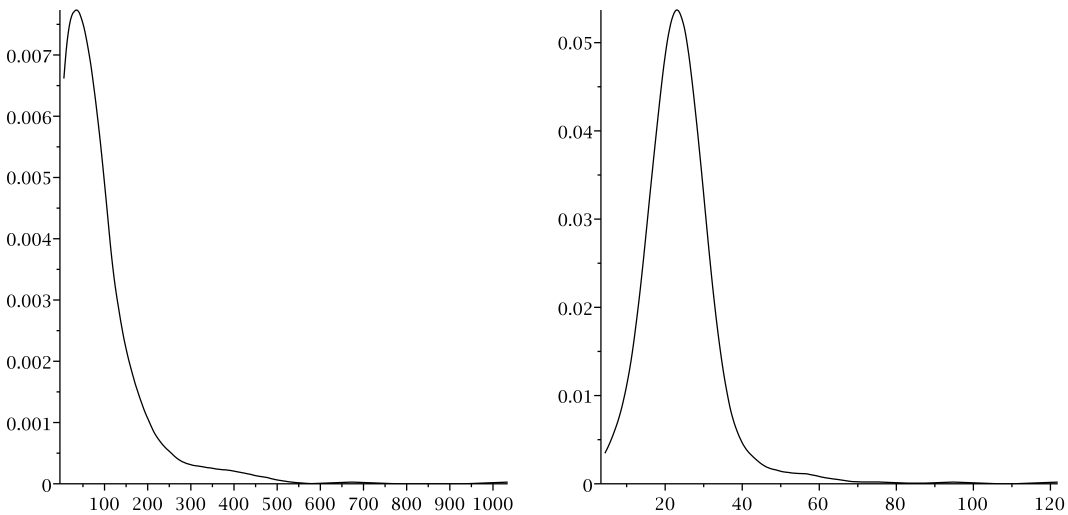
Finally, following the discussion on (10), (11), (12) in Section 3, we investigate the role of p in the conditioning of $H^{[t]}$.

Example 3 For each try, we choose $m = 10$ random polynomials with $50 \le \deg(f) \le 100$. The coefficients are randomly generated as integers between -5 and 5. The number of the terms of f is t, shown in the 3-th column of Table 3. In Table 1, *Term Range* records the range of random noise added to the evaluations of the black box polynomial; $\kappa_1({H_1}^{[t]})$ and $\kappa_1({H_2}^{[t]})$ denote the respective condition numbers of the $t \times t$ Hankel matrices ${H_1}^{[t]}$ and ${H_2}^{[t]}$. The matrices ${H_1}^{[t]}$ and ${H_2}^{[t]}$ are constructed by the approximate evaluations of the root of unity with orders corresponding to



(a) Distribution $r_{\text{low}} = \dfrac{\|H\|_1 \cdot \|H^{-1}\|_1}{\kappa_{\text{low}}(H)}$ (b) Distribution $r_{\text{up}} = \dfrac{\kappa_{\text{up}}(H)}{\|H\|_1 \cdot \|H^{-1}\|_1}$

Figure 1: Distribution of the leading principle matrix $H^{[5]}$



(a) Distribution $r_{\text{low}} = \dfrac{\|H\|_1 \cdot \|H^{-1}\|_1}{\kappa_{\text{low}}(H)}$ (b) Distribution $r_{\text{up}} = \dfrac{\kappa_{\text{up}}(H)}{\|H\|_1 \cdot \|H^{-1}\|_1}$

Figure 2: Distribution of the leading principle matrix $H^{[20]}$

the randomly chosen prime numbers p_1 and p_2, where p_1 is between 100 and 1000 and p_2 between 10^5 and 10^7.

Table 3 agrees with our discussion at the end of Section 3. When the number of terms t is fixed, the various scales of p_1 and p_2 do not seem to greatly affect the distribution of terms on the unit circle. But a larger p_2 may demand a higher precision hence cause additional computation problems.

5. CONCLUSION

By example of a randomized sparse interpolation algorithm, we have investigated a central question in algorithms with floating point arithmetic and imprecise date posed in [20]: how does one analyze the probability of success, in our case the correct determination of the discrete polynomial sparsity? Bad random choices may result in ill-conditioned intermediate structured matrices at the wrong place. Even if one can detect such ill-conditionedness, as we have via the Gohberg-Semencul formula, the distribution of such occurrences must be controlled. We do so by running multiple random execution paths next to one another, which successfully can weed out bad random choices, as we have observed experimentally. The mathematical justification requires an understanding of condition numbers of random Fourier matrices, which we have presented, and of additional instabilities that are caused by substantial noise in the data, which we have demonstrated, at least by experiment, to be con-

Ex.	*Random Noise*	t	$\kappa_1(H_1{}^{[t]})$	$\kappa_1(H_2{}^{[t]})$
1	$10^{-6} \sim 10^{-5}$	5	1.16×10^5	1.61×10^5
2	$10^{-6} \sim 10^{-5}$	8	9.54×10^5	2.91×10^6
3	$10^{-8} \sim 10^{-7}$	10	4.14×10^7	1.02×10^8
4	$10^{-6} \sim 10^{-5}$	12	1.51×10^6	2.04×10^6
5	$10^{-6} \sim 10^{-5}$	15	1.13×10^8	3.53×10^8
6	$10^{-7} \sim 10^{-6}$	12	5.24×10^7	3.92×10^7

Table 3: Condition numbers for $H^{[t]}$ constructed with roots of unity of different size prime orders p_1 and p_2.

trollable by our algorithms.

Acknowledgments

We thank anonymous referees for their helpful comments and information.

6. REFERENCES

[1] B. Beckermann. The stable computation of formal orthogonal polynomials. *Numerical Algorithms*, 11(1):1–23, 1996.

[2] B. Beckermann, S. Cabay, and G. Labahn. Fraction-free computation of matrix Pade systems. In W. Küchlin, editor, *Proc. 1997 Internat. Symp. Symbolic Algebraic Comput. (ISSAC'97)*, pages 125–132, 1997.

[3] B. Beckermann, G. H. Golub, and G. Labahn. On the numerical condition of a generalized Hankel eigenvalue problem. *Numerische Mathematik*, 106(1):41–68, 2007.

[4] B. Beckermann and G. Labahn. A fast and numerically stable euclidean-like algorithm for detecting relatively prime numerical polynomials. *Journal of Symbolic Computation*, 26(6):691–714, 1998.

[5] B. Beckermann and G. Labahn. When are two numerical polynomials relatively prime? *Journal of Symbolic Computation*, 26(6):677–689, 1998.

[6] B. Beckermann and G. Labahn. Effective computation of rational approximants and interpolants. *Reliable Computing*, 6(4):365–390, 2000.

[7] R. P. Brent, F. G. Gustavson, and D. Y. Y. Yun. Fast solution of Toeplitz systems of equations and computation of Padé approximants. *J. Algorithms*, 1:259–295, 1980.

[8] S. Cabay and R. Meleshko. A weakly stable algorithm for Padé approximants and the inversion of Hankel matrices. *SIAM Journal on Matrix Analysis and Applicatins*, 14(3):735–738, 1993.

[9] W. Gautschi. On inverse of Vandermonde and confluent Vandermonde matrices. *Numerische Mathematik*, 4:117–123, 1962.

[10] W. Gautschi. Norm estimates for inverse of Vandermonde matrices. *Numerische Mathematik*, 23:337–347, 1975.

[11] M. Giesbrecht, G. Labahn, and W. Lee. Symbolic-numeric sparse interpolation of multivariate polynomials (extended abstract). In J.-G. Dumas, editor, *ISSAC MMVI Proc. 2006 Internat. Symp. Symbolic Algebraic Comput.*, pages 116–123. ACM Press, 2006.

[12] M. Giesbrecht, G. Labahn, and W. Lee. Symbolic-numeric sparse interpolation of multivariate polynomials. *Journal of Symbolic Computation*, 44:943–959, 2009.

[13] M. Giesbrecht and D. Roche. Diversification improves interpolation. In A. Leykin, editor, *ISSAC 2011 Proc. 2011 Internat. Symp. Symbolic Algebraic Comput.* ACM Press, 2011. to appear.

[14] I. Gohberg and I. Koltracht. Efficient algorithm for Toeplitz plus Hankel matrices. *Integral Equations and Operator Theory*, 12:136–142, 1989.

[15] G. H. Golub, P. Milanfar, and J. Varah. A stable numerical method for inverting shape from moments. *SIAM Journal on Scientific Computing*, 21(4):1222–1243, 1999.

[16] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, third edition, 1996.

[17] T. Huckle. Computations with Gohberg-Semencul formulas for Toeplitz matrices. *Linear Algebra and Applications*, 271:169–198, 1998.

[18] E. Kaltofen and W. Lee. Early termination in sparse interpolation algorithms. *Journal of Symbolic Computation*, 36(3–4):365–400, 2003. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/03/KL03.pdf.

[19] E. Kaltofen, Z. Yang, and L. Zhi. Approximate greatest common divisors of several polynomials with linearly constrained coefficients and singular polynomials. In J.-G. Dumas, editor, *ISSAC MMVI Proc. 2006 Internat. Symp. Symbolic Algebraic Comput.*, pages 169–176, New York, N. Y., 2006. ACM Press. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/06/KYZ06.pdf.

[20] E. Kaltofen, Z. Yang, and L. Zhi. On probabilistic analysis of randomization in hybrid symbolic-numeric algorithms. In S. M. Watt and J. Verschelde, editors, *SNC'07 Proc. 2007 Internat. Workshop on Symbolic-Numeric Comput.*, pages 11–17, 2007. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/07/KYZ07.pdf.

[21] E. Kaltofen and G. Yuhasz. A fraction free matrix Berlekamp/Massey algorithm, Feb. 2009. Manuscript, 17 pages.

[22] S. M. Rump. Structured perturbations part I: Normwise distances. *SIAM Journal on Matrix Analysis and Applications*, 25(1):1–30, 2003.

[23] W. F. Trench. An algorithm for the inversion of finite Hankel matrices. *Journal of the Society for Industrial and Applied Mathematics*, 13(4):pp. 1102–1107, 1965.