A FAST BLOCK HANKEL SOLVER BASED ON AN INVERSION FORMULA FOR BLOCK LOEWNER MATRICES

Peter Kravanja, Marc Van Barel[(1)]

Abstract - We propose a new $\mathcal{O}(p^3 n^2)$ algorithm for solving complex $np \times np$ linear systems that have block Hankel structure, where the blocks are square matrices of size $p \times p$. Via FFTs the block Hankel system is transformed into a block Loewner system. An inversion formula enables us to calculate the inverse of the block Loewner matrix explicitely. The parameters that occur in this inversion formula are calculated by solving two rational interpolation problems on the unit circle. We have implemented our algorithm in Fortran 90. Numerical examples are included.

1. Introduction

Let n and p be positive integers. Consider a sequence $H_0, \ldots, H_{2n-2}$ in $\mathbb{C}^{p \times p}$ such that the $np \times np$ block Hankel matrix $H := [H_{k+l}]_{k,l=0}^{n-1}$ is regular. Let $b \in \mathbb{C}^{np}$. We consider the problem of computing $x := H^{-1} b$.

The $np \times np$ block exchange matrix $E := [\delta_{k,n-1-l}]_{k,l=0}^{n-1} \otimes I_p$ transforms the block Hankel system $Hx = b$ into the block Toeplitz system $T\tilde{x} = b$ where $T := HE$ and $\tilde{x} := Ex$. For a given value of p, the classical algorithms for solving a block Hankel or a block Toeplitz system exploit the structure and require less arithmetic operations, compared to $\mathcal{O}(n^3)$ for general linear systems. The so-called "fast" algorithms require $\mathcal{O}(n^2)$ arithmetic operations, while the "superfast" ones need only $\mathcal{O}(n \log^2 n)$ operations by using a divide and conquer strategy. The flow of these methods is determined by the exact singularity of the square leading principal block submatrices of H or T. The "fast" methods compute the solutions corresponding to successive nonsingular leading principal submatrices (sections). However, in finite-precision arithmetic not only singular

— Received: 16/11/1996

(1) Katholieke Universiteit Leuven, Department of Computer Science. Celestijnenlaan 200 A, B-3001 Heverlee, Belgium. E-mail: na.pkravanja@na-net.ornl.gov and marc@cs.kuleuven.ac.be

but also ill-conditioned sections should be avoided. The algorithms that have been developed for this purpose are called "look-ahead" algorithms. They look ahead from one well-conditioned section to the next one and jump over the ill-conditioned sections that lie in between. For a more detailed overview concerning scalar Hankel and Toeplitz systems, we refer the interested reader to the introduction of our paper [10] and the references cited therein. For the block Toeplitz and the block Hankel case, the reader may look at the introduction and the references of [13].

Good "look-ahead" strategies are difficult to design. Only recently, a totally different approach was considered. This approach transforms the (block) Hankel or Toeplitz matrix into a generalized Cauchy matrix. Such a matrix can be factorized using pivoting without destroying the structure. Hence, the Cauchy system can be solved without using look-ahead and the solution can be transformed back to the solution of the original system. We refer the interested reader to [5, 9, 6, 7, 8, 4] and the references cited therein. In [10] we presented such a transformation approach based on an inversion formula for Loewner matrices. In this paper, we will generalize this approach to the block case.

The outline of this paper is as follows. In Section 2 we will show how the block Hankel system $Hx = b$ can be transformed into a block Loewner system $Lx' = b'$ in $\mathcal{O}(p^2 n \log n)$ flops. An explicit formula for L^{-1} enables us to calculate x' as $L^{-1}b'$. This inversion formula for block Loewner matrices is discussed in Section 3. It involves certain $p \times p$ matrices that can be computed by solving two rational interpolation problems on the unit circle. In Section 4 we present an $\mathcal{O}(p^3 n^2)$ algorithm to solve these interpolation problems. We conclude with numerical examples in Section 5.

2. Transformation into a block Loewner system

Let $y_1, \ldots, y_n, z_1, \ldots, z_n$ be $2n$ mutually distinct complex numbers, and define $\mathbf{y} := (y_1, \ldots, y_n)$ and $\mathbf{z} := (z_1, \ldots, z_n)$. Let $\mathcal{L}(\mathbf{y}, \mathbf{z})$ be the class of $np \times np$ block matrices

$$\mathcal{L}(\mathbf{y}, \mathbf{z}) := \left\{ \left[\frac{C_k - D_l}{y_k - z_l} \right]_{k,l=1}^{n} \;\middle|\; C_1, \ldots, C_n, D_1, \ldots, D_n \in \mathbb{C}^{p \times p} \right\}.$$

The elements of $\mathcal{L}(\mathbf{y}, \mathbf{z})$ are called block Loewner matrices. They bear the name of Karl Loewner who studied them (for $p = 1$) in the context of rational interpolation and monotone matrix functions [12].

The set $\mathcal{L}(\mathbf{y}, \mathbf{z})$ is a $\mathbb{C}^{p \times p}$-module, and a submodule of the $\mathbb{C}^{p \times p}$-module of all the complex $np \times np$ block matrices having $p \times p$ blocks. Since addition of a constant matrix to all the $2n$ $p \times p$ matrices C_k, D_l leads to the same block Loewner matrix, its dimension is $2n - 1$. The set of all the complex

$np \times np$ block Hankel matrices having $p \times p$ blocks also forms a submodule of dimension $2n-1$. Block Hankel and block Loewner matrices are even more closely related. In the scalar case ($p=1$), a theorem of Fiedler [3] asserts that every Hankel matrix can be transformed into a Loewner matrix, and vice versa. In this section we will formulate this theorem for the block case and discuss our $\mathcal{O}(p^2 n \log n)$ implementation of the transformation.

First we have to deal with some preliminaries concerning Vandermonde matrices. Let $t_1, \ldots, t_n$ be n complex numbers and define $\mathbf{t} := (t_1, \ldots, t_n)$. The Vandermonde matrix with nodes $t_1, \ldots, t_n$ is given by

$$V(\mathbf{t}) = V(t_1, \ldots, t_n) := \begin{bmatrix} 1 & t_1 & \cdots & t_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & t_n & \cdots & t_n^{n-1} \end{bmatrix}.$$

Let $f_{\mathbf{t}}(z)$ be the monic polynomial of degree n that has zeros $t_1, \ldots, t_n$,

$$f_{\mathbf{t}}(z) := (z-t_1)\cdots(z-t_n),$$

and define

$$f_{\mathbf{t},j}(z) := \prod_{k \neq j}(z-t_k), \qquad j = 1, \ldots, n.$$

Note that $f_{\mathbf{t},j}(z)$ is a monic polynomial of degree $n-1$ for $j = 1, \ldots, n$. Define the $n \times n$ matrix $W(\mathbf{t})$ by the equation

$$\text{(2.1)} \quad \begin{bmatrix} f_{\mathbf{t},1}(z) \\ f_{\mathbf{t},2}(z) \\ \vdots \\ f_{\mathbf{t},n}(z) \end{bmatrix} = W(\mathbf{t}) \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{n-1} \end{bmatrix}.$$

This means that the jth row of $W(\mathbf{t})$ contains the coefficients of $f_{\mathbf{t},j}(z)$ when written in terms of the standard monomial basis $(1, z, \ldots, z^{n-1})$. Then

$$\text{(2.2)} \quad W(\mathbf{t})\,[V(\mathbf{t})]^T = \operatorname{diag}(\, f_{\mathbf{t},1}(t_1), \ldots, f_{\mathbf{t},n}(t_n)\,) =: D(\mathbf{t}).$$

The Vandermonde matrix $V(\mathbf{t})$ is regular if and only if its nodes $t_1, \ldots, t_n$ are mutually distinct. In that case (2.2) implies that $W(\mathbf{t})$ is regular, and also that

$$\text{(2.3)} \quad [V(\mathbf{t})]^{-1} = [W(\mathbf{t})]^T\,[D(\mathbf{t})]^{-1}.$$

Let $V(\mathbf{y}, \mathbf{z})$ be the $2n \times 2n$ Vandermonde matrix with nodes $y_1, \ldots, y_n$ and $z_1, \ldots, z_n$, and similarly for $W(\mathbf{y}, \mathbf{z})$.

Recall that the Kronecker product between two matrices $A = [a_{kl}]_{k,l=1}^{\alpha} \in \mathbb{C}^{\alpha \times \alpha}$ and $B \in \mathbb{C}^{\beta \times \beta}$ is defined as the block matrix

$$A \otimes B := [a_{kl} B]_{k,l=1}^{\alpha} \in \mathbb{C}^{\alpha\beta \times \alpha\beta}.$$

THEOREM 2.1. The matrix $L := [W(\mathbf{y}) \otimes I_p]\, H\, [W(\mathbf{z}) \otimes I_p]^T$ is a block Loewner matrix in $\mathcal{L}(\mathbf{y}, \mathbf{z})$ whose $p \times p$ matrices $C_1, \dots, C_n, D_1, \dots, D_n$ are given by (up to an arbitrary additive matrix $\xi \in \mathbb{C}^{p \times p}$)

$$\begin{bmatrix} C_1 \\ \vdots \\ C_n \\ D_1 \\ \vdots \\ D_n \end{bmatrix} = [W(\mathbf{y}, \mathbf{z}) \otimes I_p] \begin{bmatrix} H_0 \\ H_1 \\ \vdots \\ H_{2n-2} \\ \xi \end{bmatrix}.$$

PROOF. The scalar case ($p = 1$) has been proven by Fiedler [3, Theorem 12]. The generalization to the block case is straightforward. ∎

The Kronecker product $A \otimes B$ is regular if A and B are both regular. This implies that $W(\mathbf{y}) \otimes I_p$ and $W(\mathbf{z}) \otimes I_p$ are regular, and thus L is regular.

A judicious choice of the points $\mathbf{y}$ and $\mathbf{z}$ enables us to transform the block Hankel system $Hx = b$ into a block Loewner system $Lx' = b'$ in $\mathcal{O}(p^2 n \log n)$ flops. Let $\omega := \exp(2\pi\, i/n)$ and suppose from now on that $y_k = \omega^{k-1}$ for $k = 1, \dots, n$. That is, let $\mathbf{y} = (1, \omega, \dots, \omega^{n-1})$. Let $\zeta := \exp(\pi\, i/n)$ and suppose from now on that $z_k = \zeta y_k$ for $k = 1, \dots, n$. That is, let $\mathbf{z} = (\zeta, \zeta\omega, \dots, \zeta\omega^{n-1})$. Note that

$$\mathbf{y} = (1, \zeta^2, \dots, \zeta^{2n-2}) \qquad \text{and} \qquad \mathbf{z} = (\zeta, \zeta^3, \dots, \zeta^{2n-1}).$$

Let Ω_n be the $n \times n$ Fourier matrix,

$$(2.4)\quad \Omega_n := \frac{1}{\sqrt{n}} V(1, \omega, \dots, \omega^{n-1}).$$

Matrix-vector products involving Ω_n (Ω_n^H) amount to a(n) (inverse) discrete Fourier transform (DFT) and can be evaluated via the celebrated (inverse) fast Fourier transform (FFT) in $\mathcal{O}(n \log n)$ flops [11]. Finally, let $D_{n,\omega}$ and $D_{n,\zeta}$ be the $n \times n$ diagonal matrices

$$D_{n,\omega} := \operatorname{diag}(1, \omega, \dots, \omega^{n-1}) \qquad \text{and} \qquad D_{n,\zeta} := \operatorname{diag}(1, \zeta, \dots, \zeta^{n-1}).$$

THEOREM 2.2. The solution to the block Hankel system $Hx = b$ is given by

$$x = [\, \sqrt{n}\, \zeta^{n-1} \overline{D_{n,\zeta}}\, \Omega_n^H\, \overline{D_{n,\omega}} \otimes I_p \,] x'$$

where $x' := L^{-1} b'$ with $b' := [\, \sqrt{n}\, \overline{D_{n,\omega} \Omega_n} \otimes I_p \,] b$.

PROOF. Theorem 2.1 implies that $x = H^{-1} b$ is given by $x = [W(\mathbf{z}) \otimes I_p]^T x'$ where $x' := L^{-1} b'$ with $b' := [W(\mathbf{y}) \otimes I_p] b$. In the proof of [10, Theorem 2.2] it is shown that

$$W(\mathbf{y}) = \sqrt{n}\, \overline{D_{n,\omega} \Omega_n} \qquad \text{and} \qquad [W(\mathbf{z})]^T = \sqrt{n}\, \zeta^{n-1} \overline{D_{n,\zeta}}\, \Omega_n^H\, \overline{D_{n,\omega}}.$$

As $[W(\mathbf{z}) \otimes I_p]^T = [W(\mathbf{z})]^T \otimes I_p$, the result follows. ■

It follows that

$$\begin{bmatrix} x_r \\ x_{p+r} \\ \vdots \\ x_{(n-1)p+r} \end{bmatrix} = \sqrt{n}\, \zeta^{n-1} \overline{D_{n,\zeta}}\, \Omega_n^H\, \overline{D_{n,\omega}} \begin{bmatrix} x'_r \\ x'_{p+r} \\ \vdots \\ x'_{(n-1)p+r} \end{bmatrix}$$

and

$$\begin{bmatrix} b'_r \\ b'_{p+r} \\ \vdots \\ b'_{(n-1)p+r} \end{bmatrix} = \sqrt{n}\, \overline{D_{n,\omega} \Omega_n} \begin{bmatrix} b_r \\ b_{p+r} \\ \vdots \\ b_{(n-1)p+r} \end{bmatrix}$$

for $r = 1, \ldots, p$. Thus the transformations $b \mapsto b'$ and $x' \mapsto x$ can be done in $\mathcal{O}(pn \log n)$ floating point operations.

Let P be the permutation matrix defined by

$$P \begin{bmatrix} y_1 \\ \vdots \\ y_n \\ z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} 1 \\ \zeta \\ \vdots \\ \zeta^{2n-1} \end{bmatrix},$$

let Ω_{2n} be the $2n \times 2n$ Fourier matrix,

$$\Omega_{2n} := \frac{1}{\sqrt{2n}} V(1, \zeta, \ldots, \zeta^{2n-1}),$$

and let $D_{2n,\zeta} := \operatorname{diag}(1, \zeta, \ldots, \zeta^{2n-1})$.

Theorem 2.3. The $p \times p$ matrices $C_1, \ldots, C_n, D_1, \ldots, D_n$ of the block Loewner matrix L are given by (up to an arbitrary additive matrix $\xi \in \mathbb{C}^{p \times p}$)

$$\begin{bmatrix} C_1 \\ \vdots \\ C_n \\ D_1 \\ \vdots \\ D_n \end{bmatrix} = [\, \sqrt{2n}\, P^T \overline{D_{2n,\zeta}\, \Omega_{2n}} \otimes I_p\,] \begin{bmatrix} H_0 \\ H_1 \\ \vdots \\ H_{2n-2} \\ \xi \end{bmatrix}.$$

Proof. In [10, Theorem 2.3] it is shown that $W(\mathbf{y}, \mathbf{z}) = \sqrt{2n}\, P^T \overline{D_{2n,\zeta}\, \Omega_{2n}}$. The result then follows immediately from Theorem 2.1. ■

The previous theorem implies that

$$\begin{bmatrix} (C_1)_{k,l} \\ \vdots \\ (C_n)_{k,l} \\ (D_1)_{k,l} \\ \vdots \\ (D_n)_{k,l} \end{bmatrix} = \sqrt{2n}\, P^T \overline{D_{2n,\zeta}\,\Omega_{2n}} \begin{bmatrix} (H_0)_{k,l} \\ (H_1)_{k,l} \\ \vdots \\ (H_{2n-2})_{k,l} \\ (\xi)_{k,l} \end{bmatrix}$$

for $k, l = 1, \ldots, p$. Thus H can be transformed into L in $\mathcal{O}(p^2 n \log n)$ flops.

Note. As we already mentioned in the proof of Theorem 2.2, in [10, Theorem 2.2] it is shown that

$$W(\mathbf{y}) = \sqrt{n}\,\overline{D_{n,\omega}\Omega_n} \qquad \text{and} \qquad [W(\mathbf{z})]^T = \sqrt{n}\,\zeta^{n-1}\overline{D_{n,\zeta}}\,\Omega_n^H\,\overline{D_{n,\omega}}.$$

Thus $W(\mathbf{y})/\sqrt{n}$ and $[W(\mathbf{z})]^T/\sqrt{n}$ are unitary. The Kronecker product of two unitary matrices is a unitary matrix. As

$$L = n\,[\frac{1}{\sqrt{n}}W(\mathbf{y}) \otimes I_p]\,H\,[\frac{1}{\sqrt{n}}[W(\mathbf{z})]^T \otimes I_p],$$

it follows that $\|L\|_2 = n\|H\|_2$ and $\|L^{-1}\|_2 = n^{-1}\|H^{-1}\|_2$ and thus $\kappa_2(L) = \kappa_2(H)$. In other words, the spectral condition number of H is left unchanged.

3. An inversion formula for block Loewner matrices

THEOREM 3.1. Let P_k, U_k, $\tilde{P}_k$ and $\tilde{U}_k \in \mathbb{C}^{p\times p}$ for $k = 1, \ldots, n$ be defined by the equations

$$\begin{aligned} [\,P_1 \;\; \cdots \;\; P_n\,]\,L &= [\,I_p \;\; \cdots \;\; I_p\,] \\ [\,U_1 \;\; \cdots \;\; U_n\,]\,L &= [\,D_1 \;\; \cdots \;\; D_n\,] \\ L\begin{bmatrix} \tilde{P}_1 \\ \vdots \\ \tilde{P}_n \end{bmatrix} &= \begin{bmatrix} I_p \\ \vdots \\ I_p \end{bmatrix} \\ L\begin{bmatrix} \tilde{U}_1 \\ \vdots \\ \tilde{U}_n \end{bmatrix} &= \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix}. \end{aligned}$$

Then the inverse of L is given by

$$L^{-1} = \Big[\,\frac{\tilde{U}_k\,P_l - \tilde{P}_k\,U_l}{y_l - z_k}\,\Big]_{k,l=1}^{n}. \tag{3.1}$$

PROOF. See Van Barel and Vavřín [15, Theorem 2.2]. ∎

The $p \times p$ matrices that appear in the inversion formula (3.1) can be computed by solving the following linearized rational interpolation problems.

THEOREM 3.2. Define

$$\begin{aligned} U(z) &:= f_{\mathbf{y}}(z)I_p - \sum_{k=1}^{n} f_{\mathbf{y},k}(z)U_k \\ V(z) &:= -\sum_{k=1}^{n} f_{\mathbf{y},k}(z)U_k C_k \\ \tilde{U}(z) &:= f_{\mathbf{z}}(z)I_p - \sum_{k=1}^{n} f_{\mathbf{z},k}(z)\tilde{U}_k \\ \tilde{V}(z) &:= -\sum_{k=1}^{n} f_{\mathbf{z},k}(z)D_k\tilde{U}_k. \end{aligned}$$

Then the pair $(V(z), U(z))$ is the only pair such that

- $U(z) \in \mathbb{C}[z]^{p\times p}$ and $\deg U(z) = n$, $V(z) \in \mathbb{C}[z]^{p\times p}$ and $\deg V(z) < n$, $U(z)$ is monic
- $V(y_k) = U(y_k)C_k$ and $V(z_k) = U(z_k)D_k$ for $k = 1, \ldots, n$.

The pair $(\tilde{V}(z), \tilde{U}(z))$ is the only pair such that

- $\tilde{U}(z) \in \mathbb{C}[z]^{p\times p}$ and $\deg \tilde{U}(z) = n$, $\tilde{V}(z) \in \mathbb{C}[z]^{p\times p}$ and $\deg \tilde{V}(z) < n$, $\tilde{U}(z)$ is monic
- $\tilde{V}(y_k) = C_k\tilde{U}(y_k)$ and $\tilde{V}(z_k) = D_k\tilde{U}(z_k)$ for $k = 1, \ldots, n$.

Moreover, these pairs represent the same matrix rational function,

$$[U(z)]^{-1}V(z) \equiv \tilde{V}(z)[\tilde{U}(z)]^{-1}.$$

PROOF. See Van Barel and Vavřín [15, Theorem 3.1]. ∎

Observe that $U(z)$ is invertible a.e. Indeed, the fact that $U(z)$ is monic implies that $\det U(z) = z^{pn} +$ lower order terms, and thus $\det U(z) \neq 0$ a.e. The same holds for $\tilde{U}(z)$.

THEOREM 3.3. Define

$$\begin{aligned} P(z) &:= \sum_{k=1}^{n} f_{\mathbf{y},k}(z)P_k \\ Q(z) &:= f_{\mathbf{y}}(z)I_p + \sum_{k=1}^{n} f_{\mathbf{y},k}(z)P_k C_k \\ \tilde{P}(z) &:= \sum_{k=1}^{n} f_{\mathbf{z},k}(z)\tilde{P}_k \\ \tilde{Q}(z) &:= f_{\mathbf{z}}(z)I_p + \sum_{k=1}^{n} f_{\mathbf{z},k}(z)D_k\tilde{P}_k. \end{aligned}$$

Then the pair $(Q(z), P(z))$ is the only pair such that

- $Q(z) \in \mathbb{C}[z]^{p\times p}$ and $\deg Q(z) = n$, $P(z) \in \mathbb{C}[z]^{p\times p}$ and $\deg P(z) < n$, $Q(z)$ is monic
- $Q(y_k) = P(y_k)C_k$ and $Q(z_k) = P(z_k)D_k$ for $k = 1, \dots, n$.

The pair $(\tilde{Q}(z), \tilde{P}(z))$ is the only pair such that

- $\tilde{Q}(z) \in \mathbb{C}[z]^{p\times p}$ and $\deg \tilde{Q}(z) = n$, $\tilde{P}(z) \in \mathbb{C}[z]^{p\times p}$ and $\deg \tilde{P}(z) < n$, $\tilde{Q}(z)$ is monic
- $\tilde{Q}(y_k) = C_k\tilde{P}(y_k)$ and $\tilde{Q}(z_k) = D_k\tilde{P}(z_k)$ for $k = 1, \dots, n$.

Moreover, these pairs represent the same matrix rational function,

$$[Q(z)]^{-1}P(z) \equiv \tilde{P}(z)[\tilde{Q}(z)]^{-1}.$$

PROOF. See Van Barel and Vavřín [15, Theorem 3.2]. ■

Observe that $Q(z)$ and $\tilde{Q}(z)$ are monic and thus invertible a.e. Note that if $U(y_k)$ and $U(z_k)$, $k = 1, \dots, n$, are regular, the matrix rational function $[U(z)]^{-1}V(z)$ satisfies the proper rational interpolation conditions

$$[U(y_k)]^{-1}V(y_k) = C_k \quad \text{and} \quad [U(z_k)]^{-1}V(z_k) = D_k$$

for $k = 1, \dots, n$. Similarly, if $P(y_k)$ and $P(z_k)$, $k = 1, \dots, n$, are regular, the matrix rational function $[P(z)]^{-1}Q(z)$ satisfies the proper rational interpolation conditions

$$[P(y_k)]^{-1}Q(y_k) = C_k \quad \text{and} \quad [P(z_k)]^{-1}Q(z_k) = D_k$$

for $k = 1, \dots, n$. The matrix rational functions $[U(z)]^{-1}V(z)$ and $[P(z)]^{-1}Q(z)$ are different because their degree structure is different. The same conclusions hold for $\tilde{V}(z)[\tilde{U}(z)]^{-1}$ and $\tilde{Q}(z)[\tilde{P}(z)]^{-1}$.

In the next section we will present an algorithm that solves these linearized rational interpolation problems. An easy calculation reveals the following connections with the $p \times p$ matrices that appear in the inversion formula (3.1):

$$\begin{bmatrix} U_1 \\ \vdots \\ U_n \end{bmatrix} = -\frac{1}{n}[D_{n,\omega} \otimes I_p] \begin{bmatrix} U(y_1) \\ \vdots \\ U(y_n) \end{bmatrix}, \quad \begin{bmatrix} P_1 \\ \vdots \\ P_n \end{bmatrix} = \frac{1}{n}[D_{n,\omega} \otimes I_p] \begin{bmatrix} P(y_1) \\ \vdots \\ P(y_n) \end{bmatrix},$$

$$\begin{bmatrix} \tilde{U}_1 \\ \vdots \\ \tilde{U}_n \end{bmatrix} = -\frac{1}{n}\overline{\zeta^{n-1}}\,[D_{n,\omega} \otimes I_p] \begin{bmatrix} \tilde{U}(z_1) \\ \vdots \\ \tilde{U}(z_n) \end{bmatrix}, \quad \begin{bmatrix} \tilde{P}_1 \\ \vdots \\ \tilde{P}_n \end{bmatrix} = \frac{1}{n}\overline{\zeta^{n-1}}\,[D_{n,\omega} \otimes I_p] \begin{bmatrix} \tilde{P}(z_1) \\ \vdots \\ \tilde{P}(z_n) \end{bmatrix}.$$

As the points $\mathbf{y}$ are the nth roots of unity and $\mathbf{z} = \zeta\mathbf{y}$, the right hand sides of these formulae can be evaluated in $\mathcal{O}(p^2 n \log n)$ flops. The product $x' = L^{-1}b'$

can then be calculated in $4np^2 + \mathcal{O}(pn\log n)$ flops. Indeed, the matrix L^{-1} can be written as

$$(3.2)\quad L^{-1} = \operatorname{diag}(\tilde{P}_1,\dots,\tilde{P}_n)\,C\,\operatorname{diag}(U_1,\dots,U_n) - \operatorname{diag}(\tilde{U}_1,\dots,\tilde{U}_n)\,C\,\operatorname{diag}(P_1,\dots,P_n)$$

if C is given by the $np \times np$ block Cauchy matrix

$$C := \Big[\,\frac{I_p}{z_k - y_l}\,\Big]_{k,l=1}^{n}.$$

Note that $C = C_1 \otimes I_p$ if C_1 is the corresponding $n \times n$ Cauchy matrix

$$C_1 := \Big[\,\frac{1}{z_k - y_l}\,\Big]_{k,l=1}^{n}.$$

The following theorem implies that the product of C_1 with a vector in $\mathbb{C}^n$ can be evaluated in $\mathcal{O}(n\log n)$ flops.

THEOREM 3.4. Let $v = [\,v_1 \quad \cdots \quad v_n\,]^T$ be a vector in $\mathbb{C}^n$. Then

$$C_1 v = -\frac{n}{2}\,\Omega_n\,D_{n,\zeta}\,\Omega_n^H\,\overline{D_{n,\omega}}\,v.$$

PROOF. See Kravanja and Van Barel [10, Theorem 3.4]. ■

Thus, because of (3.2), the product $L^{-1}b'$ can indeed be evaluated in $4np^2 + \mathcal{O}(pn\log n)$ floating point operations.

4. Solving the linearized rational interpolation problems

Let $s_k := y_k$ and $s_{n+k} := z_k$ for $k = 1,\dots,n$. Define the block row vectors $f_{\lambda,1},\dots,f_{\lambda,2n} \in \mathbb{C}^{p\times 2p}$ as

$$f_{\lambda,k} := [\,I_p \quad -C_k\,], \quad f_{\lambda,n+k} := [\,I_p \quad -D_k\,], \qquad k = 1,\dots,n.$$

Consider the interpolation problem

$$(4.1)\quad f_{\lambda,k}B_\lambda(s_k) = [\,O_p \quad O_p\,], \qquad k = 1,\dots,2n,$$

where

$$B_\lambda(z) = \begin{bmatrix} N_l(z) & N_r(z) \\ D_l(z) & D_r(z) \end{bmatrix} \in \mathbb{C}[z]^{2p\times 2p}$$

where $\deg N_l(z) = n$, $\deg D_l(z) \le n-1$, $\deg N_r(z) \le n-1$ and $\deg D_r(z) = n$, and $N_l(z)$ as well as $D_r(z)$ monic, i.e., where $B_\lambda(z)$ is a monic $2p \times 2p$ block matrix polynomial of degree n. Theorem 3.2 and 3.3 assert that this problem has a unique solution given by

$$B_\lambda^\star(z) := \begin{bmatrix} \tilde{Q}(z) & \tilde{V}(z) \\ \tilde{P}(z) & \tilde{U}(z) \end{bmatrix}.$$

Similarly, define the block column vectors $f_{\rho,1}, \ldots, f_{\rho,2n} \in \mathbb{C}^{2p \times p}$ as

$$f_{\rho,k} := \begin{bmatrix} I_p \\ -C_k \end{bmatrix}, \quad f_{\rho,n+k} := \begin{bmatrix} I_p \\ -D_k \end{bmatrix}, \qquad k = 1, \ldots, n,$$

and consider the interpolation problem

$$B_\rho(s_k) f_{\rho,k} = \begin{bmatrix} O_p \\ O_p \end{bmatrix}, \qquad k = 1, \ldots, 2n, \tag{4.2}$$

where

$$B_\rho(z) = \begin{bmatrix} N_u(z) & D_u(z) \\ N_l(z) & D_l(z) \end{bmatrix} \in \mathbb{C}[z]^{2p \times 2p}$$

where $\deg N_u(z) = n$, $\deg D_u(z) \le n-1$, $\deg N_l(z) \le n-1$ and $\deg D_l(z) = n$, and $N_u(z)$ as well as $D_l(z)$ monic, i.e., where $B_\rho(z)$ is a monic $2p \times 2p$ block matrix polynomial of degree n. Theorem 3.2 and 3.3 assert that this problem has a unique solution given by

$$B_\rho^\star(z) := \begin{bmatrix} Q(z) & P(z) \\ V(z) & U(z) \end{bmatrix}.$$

Algorithm 4.1 calculates a $2p \times 2p$ matrix polynomial $B_\lambda(z)$ of degree n that satisfies (4.1) and whose highest degree coefficient $A_\lambda \in \mathbb{C}^{2p \times 2p}$ is regular. This implies that $B_\lambda(z) \equiv B_\lambda^\star(z) A_\lambda$. Hence, to obtain $B_\lambda^\star(z)$ we have to multiply $B_\lambda(z)$ to the right by the inverse of A_λ. An analogous algorithm can be formulated that calculates a $2p \times 2p$ matrix polynomial $B_\rho(z)$ of degree n that satisfies (4.2) and whose highest degree coefficient $A_\rho \in \mathbb{C}^{2p \times 2p}$ is regular, which implies that $B_\rho(z) \equiv A_\rho B_\rho^\star(z)$. We leave this to the reader. We will show below that $\det A_\lambda = \det A_\rho = 1$. Note that only the second block row of $B_\lambda^\star(z)$ and the second block column of $B_\rho^\star(z)$ are needed to compute the inversion parameters.

ALGORITHM 4.1.

input $n, p;\ y_1, \ldots, y_n, z_1, \ldots, z_n;\ C_1, \ldots, C_n, D_1, \ldots, D_n$
output $B_\lambda(z)$, $\alpha > 0$
initialization

$$e \leftarrow \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^p; \quad B_\lambda(z) \leftarrow \begin{bmatrix} I_p & O_p \\ O_p & I_p \end{bmatrix}; \quad s \leftarrow \begin{bmatrix} y_1 \\ \vdots \\ y_n \\ z_1 \\ \vdots \\ z_n \end{bmatrix} \otimes e$$

$$\text{Residuals} \leftarrow \begin{bmatrix} I_p & -C_1 \\ \vdots & \vdots \\ I_p & -C_n \\ I_p & -D_1 \\ \vdots & \vdots \\ I_p & -D_n \end{bmatrix} ; \mathcal{C} \leftarrow \{1, \ldots, 2p\}$$

comment

Note that $s \in \mathbb{C}^{2pn}$ and Residuals $\in \mathbb{C}^{2pn \times 2p}$.

1. $\alpha \leftarrow \max\limits_{\substack{1 \le k \le 2pn \\ p+1 \le l \le 2p}} \{\max\{|\operatorname{Re} \text{Residuals}(k,l)|, |\operatorname{Im} \text{Residuals}(k,l)|\}\}$

2. $\text{Residuals}(1:2pn, p+1:2p) \leftarrow \text{Residuals}(1:2pn, p+1:2p)/\alpha$

for $j = 1 : 2pn$

3. Determine $\text{piv} \in \{j, \ldots, 2pn\}$ and $\text{col} \in \mathcal{C}$ such that
$\max\{|\operatorname{Re} \text{Residuals}(\text{piv}, \text{col})|, |\operatorname{Im} \text{Residuals}(\text{piv}, \text{col})|\}$
$= \max\limits_{\substack{j \le k \le 2pn \\ l \in \mathcal{C}}} \{\max\{|\operatorname{Re} \text{Residuals}(k,l)|, |\operatorname{Im} \text{Residuals}(k,l)|\}\}.$

4. $s(j) \leftrightarrow s(\text{piv})$; $\text{Residuals}(j,:) \leftrightarrow \text{Residuals}(\text{piv},:)$

5. **for** l **in** $\{1, \ldots, 2p\} \setminus \{\text{col}\}$

   5.1 $\mu(l) \leftarrow \text{Residuals}(j,l)/\text{Residuals}(j,\text{col})$

   end for

   5.2 $$\text{B}_\lambda(\text{z}) \leftarrow \text{B}_\lambda(\text{z}) \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ -\mu(1) & \cdots & -\mu(\text{col}-1) & z - s(j) & -\mu(\text{col}+1) & \cdots & -\mu(2p) \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

   for $k = j+1 : 2pn$

   for l **in** $\{1, \ldots, 2p\} \setminus \{\text{col}\}$

   5.3a $\text{Residuals}(k,l) \leftarrow \text{Residuals}(k,l) - \mu(l)\text{Residuals}(k,\text{col})$

   end for

   5.3b $\text{Residuals}(k,\text{col}) \leftarrow (s(k) - s(j))\text{Residuals}(k,\text{col})$

   end for

6. $\mathcal{C} \leftarrow \mathcal{C} \setminus \{\text{col}\}$; **if** $\mathcal{C} == \emptyset$ **then** $\mathcal{C} \leftarrow \{1, \ldots, 2p\}$

end for

Notes

We state the algorithm without proof as one can easily verify that it works correctly. It generalizes the algorithm that appears in [10] to the block case.

The scaling in step 2 is done for reasons of numerical stability. If we divide the interpolation data $C_1, \ldots, C_n, D_1, \ldots, D_n$ by α, then we do not solve the original interpolation problem, of course, but rather the one that is connected with the block Loewner matrix L/α. By applying the inversion formula we now obtain the matrix αL^{-1} and thus $x' = L^{-1}b'$ is given by $x' = \tilde{x}'/\alpha$ where $\tilde{x}' := (\alpha L^{-1})b'$.

Step 5.2 of Algorithm 4.1 implies that $\det B_\lambda(z) \leftarrow \det B_\lambda(z)(z - s(j))$ for $j = 1 : 2pn$. It follows that

$$\det B_\lambda(z) \equiv (z-y_1)^p \cdots (z-y_n)^p (z-z_1)^p \cdots (z-z_n)^p,$$

a monic polynomial of degree $2pn$. As $\det B_\lambda(z) \equiv \det B_\lambda^\star(z) \det A_\lambda$ and $\det B_\lambda^\star(z)$ is also a monic polynomial of degree $2pn$, we may conclude that $\det A_\lambda = 1$. A similar argument shows that $\det A_\rho = 1$.

To calculate the inversion parameters we first have to evaluate the $p \times p$ matrix polynomials that appear in the second block row of $B_\lambda^\star(z)$ in the points $z_1, \ldots, z_n$, and those that appear in the second block column of $B_\rho^\star(z)$ in the points $y_1, \ldots, y_n$. This can be done via FFTs in $4p^2\mathcal{O}(n \log n)$ flops.

What is the arithmetic complexity of Algorithm 4.1? As we have already said, only the second block row of $B_\lambda^\star(z)$ needs to be computed. This takes $2p^3n^2 + \mathcal{O}(n)$ complex multiplications and $2p^3n^2 + \mathcal{O}(n)$ complex additions. The updating of the residuals costs $4p^3n^2 + \mathcal{O}(n)$ complex multiplications and $4p^3n^2 + \mathcal{O}(n)$ complex additions. The overall cost to compute the second block row of $B_\lambda^\star(z)$ is therefore $6p^3n^2 + \mathcal{O}(n)$ complex multiplications and $6p^3n^2 + \mathcal{O}(n)$ complex additions. Similarly we need $6p^3n^2 + \mathcal{O}(n)$ complex multiplications and $6p^3n^2 + \mathcal{O}(n)$ complex additions to compute the second block column of $B_\rho^\star(z)$.

The multipliers $\mu(l)$ that have to be computed in step 5.1 are well defined. Indeed, consider the set $\mathcal{S}$ of all the column vector polynomials $w(z) \in \mathbb{C}[z]^{2p\times 1}$ that satisfy the interpolation conditions

$$(4.3) \quad f_{\lambda,k} w(s_k) = O_{p\times 1}, \qquad k = 1, \ldots, 2n.$$

If after the execution of step $j < 2pn$ all the residuals would be equal to zero, then the interpolation problem (4.3) would have a solution of degree $< n$. This is impossible, as we will now show. The set $\mathcal{S}$ forms a submodule of the $\mathbb{C}[z]$-module $\mathbb{C}[z]^{2p\times 1}$. A basis for $\mathcal{S}$ always consists of exactly $2p$ elements, i.e., the dimension of $\mathcal{S}$ is equal to $2p$ [14, Theorem 3.1]. Let $\{B_k(z)\}_{k=1}^{2p}$ be a basis for $\mathcal{S}$. Then every element $w(z) \in \mathcal{S}$ can be written in a unique way as $w(z) = \sum_{k=1}^{2p} \alpha_k(z) B_k(z)$ with $\alpha_k(z) \in \mathbb{C}[z]$ for $k = 1, \ldots, 2p$. The matrix polynomial $B(z) := [\, B_1(z) \; B_2(z) \; \cdots \; B_{2p}(z) \,] \in \mathbb{C}[z]^{2p\times 2p}$ is called a *basis matrix*. Basis matrices can be characterized as follows.

THEOREM 4.1. A matrix polynomial $C(z) = [\, C_1(z)\; C_2(z)\; \cdots\; C_{2p}(z)\,] \in \mathbb{C}[z]^{2p\times 2p}$ is a basis matrix if and only if $C_k(z) \in \mathcal{S}$ for $k = 1, \ldots, 2p$ and $\deg \det C(z) = 2pn$.

PROOF. This follows immediately from [14, Theorem 4.1]. ■

Note that $B^\star_\lambda(z)$ is a basis matrix.

A matrix polynomial is called *column reduced* if the highest degree coefficients of its column vector polynomials are linearly independent. Every basis matrix can be transformed into a column reduced basis matrix [14, p. 455]. Note that $B^\star_\lambda(z)$ is column reduced.

THEOREM 4.2. Let $\delta_k := \deg B_k(z)$ for $k = 1, \ldots, 2p$. If $B(z)$ is column reduced, then every element $w(z) \in \mathcal{S}$ having degree $\leq \delta$ can be written in a unique way as $w(z) = \sum_{k=1}^{2p} \alpha_k(z) B_k(z)$ with $\alpha_k(z) \in \mathbb{C}[z]$ and $\deg \alpha_k(z) \leq \delta - \delta_k$ for $k = 1, \ldots, 2p$.

PROOF. See [14, Theorem 3.2]. ■

COROLLARY 4.1. The interpolation problem (4.3) has no nontrivial solution of degree $< n$.

PROOF. $B^\star_\lambda(z)$ is a column reduced basis matrix. Its column degrees are equal to n. Suppose that $\delta < n$. Then $\deg \alpha_k(z) < 0$ and thus $\alpha_k(z) \equiv 0$ for $k = 1, \ldots, 2p$. This implies that $w(z) \equiv O_{2p\times 1}$. ■

5. Numerical examples

We have implemented our solver in Fortran 90 and in Matlab. In the Fortran version the FFTs are calculated via FFTPACK. The Fortran 90 package is available at URL **`http://www.cs.kuleuven.ac.be/~marc/hankel/`**. The Matlab m-files can be obtained via anonymous ftp to **`ftp.mathworks.com`**. Look for the files **`block_hsolve.m`** and **`block_ratint.m`** in the directory **`/pub/contrib/linalg`**. The corresponding m-files for the case $p = 1$ are **`hsolve.m`** and **`ratint.m`**.

A matrix-vector product involving an $n \times n$ Hankel matrix amounts to a convolution of two vectors or, equivalently, the product of two polynomials, and can thus be calculated via FFTs in $\mathcal{O}(n \log n)$ flops [2]. By writing an approximation $\hat{x}$ to $x = H^{-1}b$ as

$$\hat{x} = \left[\begin{array}{c} \hat{x}_1 \\ 0 \\ \vdots \\ 0 \\ \hline \vdots \\ \hline \hat{x}_{(n-1)p+1} \\ 0 \\ \vdots \\ 0 \end{array}\right] + \left[\begin{array}{c} 0 \\ \hat{x}_2 \\ 0 \\ \vdots \\ 0 \\ \hline \vdots \\ \hline 0 \\ \hat{x}_{(n-1)p+2} \\ 0 \\ \vdots \\ 0 \end{array}\right] + \cdots + \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ \hat{x}_p \\ \hline \vdots \\ \hline 0 \\ \vdots \\ 0 \\ \hat{x}_{np} \end{array}\right],$$

it follows that the residue $r := b - H\hat{x}$ can be calculated in $p^2 n + \mathcal{O}(p^2 n \log n)$ flops. This implies that improving an approximation for x iteratively does not add to the $\mathcal{O}(p^3 n^2)$ complexity of our algorithm.

In the following examples we have used the Fortran 90 package. The calculations have been done by an IBM SP2 with machine epsilon $\epsilon \approx 0.12\,10^{-6}$ in single precision and $\epsilon \approx 0.22\,10^{-15}$ in double precision.

Example 1. We considered single precision $np \times np$ real block Hankel matrices $H_{n,p}$ having $p \times p$ blocks whose entries were random uniformly distributed in $[-1, 1]$. The right hand sides $b_{n,p} \in \mathbb{R}^{np}$ were calculated in double precision such that

$$x_{n,p} := H_{n,p}^{-1} b_{n,p} = [\ 1 \ \ \cdots \ \ 1\ ]^T.$$

We fix the block size $p = 3$ and let $n = 1(1)100$. Figure 1 shows the results obtained by our algorithm (before and after three steps of iterative improvement in which the residue was calculated in double precision) and the results obtained via Gaussian elimination with partial pivoting (GEPP) using the LAPACK routines CGETRF and CGETRS [1]. We have plotted $-\log_{10}(\|\hat{x}_{n,p} - x_{n,p}\|_\infty / \|x_{n,p}\|_\infty)$. Also shown is $\log_{10} \kappa_\infty(H_{n,p})$. The condition number was calculated via LAPACKs routine CGECON.

Example 2. We considered the block Hankel matrix

$$H_{2n} := \begin{bmatrix} 1 & 2 & \cdots & n \\ 2 & & \iddots & 0 \\ \vdots & \iddots & \iddots & \vdots \\ n & 0 & \cdots & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$$

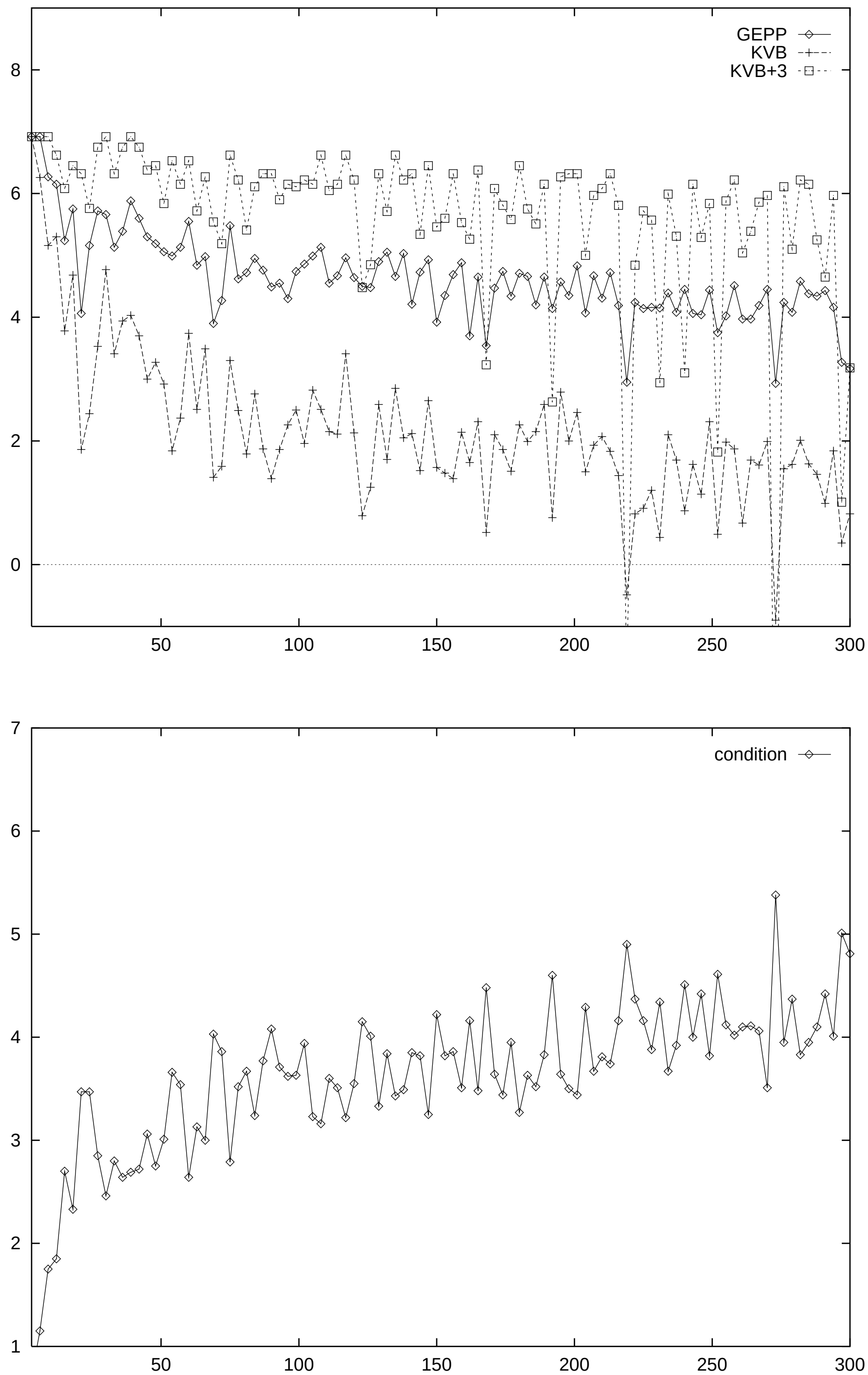

Figure 1: $-\log_{10} \dfrac{\|\hat{x}_{n,p} - x_{n,p}\|_\infty}{\|x_{n,p}\|_\infty}$ and $\log_{10} \kappa_\infty(H_{n,p})$ versus np (Example 1)

for $n = 1(1)500$. We defined the right hand side $b_{2n} = [\ b_1^{(2n)} \ \cdots \ b_{2n}^{(2n)}\ ]^T$ as

$$b_{2k-1}^{(2n)} = 0, \qquad b_{2k}^{(2n)} = n(n+1) - k(k-1), \qquad k = 1, \cdots, n.$$

Then the solution to $H_{2n}x_{2n} = b_{2n}$ is given by $x_{2n} = [\ 1 \ \cdots \ 1\ ]^T$ as one can easily verify. Figure 2 shows the results obtained by our algorithm (before and after one step of iterative improvement). The calculations were done in double precision. Also shown is $\log_{10} \kappa_\infty(H_{2n})$. The condition number was calculated via LAPACKs routine CGECON.

These examples indicate that the algorithm is weakly stable, i.e., if the block Hankel matrix is not too ill-conditioned, then the method combined with one or more steps of iterative improvement calculates the solution with high accuracy.

REFERENCES

[1] E. ANDERSON, Z. BAI, C. BISCHOF, J. DEMMEL, J. DONGARRA, J. DU CROZ, A. GREENBAUM, S. HAMMARLING, A. MCKENNEY, S. OSTROUCHOV, D. SORENSEN, LAPACK Users' Guide, SIAM, 1994.

[2] D. BINI, V. PAN, Polynomial and Matrix Computations, volume 1: Fundamental Algorithms, Birkhäuser, 1994.

[3] M. FIEDLER, *Hankel and Loewner matrices,* Linear Algebra Appl. **58**, (1984) 75–95.

[4] K. A. GALLIVAN, S. THIRUMALAI, P. VAN DOOREN, V. VERMAUT, *High performance algorithms for Toeplitz and block Toeplitz matrices,* Linear Algebra Appl. **241–243**, (1996).

[5] I. GOHBERG, T. KAILATH, V. OLSHEVSKY, *Fast Gaussian elimination with partial pivoting for matrices with displacement structure,* Math. Comput. **64 (212)**, (1995) 1557–1576.

[6] G. HEINIG, *Transformation approaches for fast and stable solution of Toeplitz systems and polynomial equations.* In Proceedings of the International Workshop "Recent Advances in Applied Mathematics," pages 223–238, State of Kuwait, May 4–7 1996.

[7] G. HEINIG, A. BOJANCZYK, *Transformation techniques for Toeplitz and Toeplitz-plus-Hankel matrices: I. Transformations,* Submitted to Linear Algebra Appl.

[8] G. HEINIG, A. BOJANCZYK, *Transformation techniques for Toeplitz and Toeplitz-plus-Hankel matrices: II. Algorithms,* Submitted to Linear Algebra Appl.

[9] T. KAILATH, A. H. SAYED, *Displacement structure: theory and applications,* SIAM Review **37**, (1995) 297–386.

[10] P. KRAVANJA, M. VAN BAREL, *A fast Hankel solver based on an inversion formula for Loewner matrices.* Submitted to Linear Algebra Appl.

[11] C. VAN LOAN, Computational Frameworks for the Fast Fourier Transform, volume 10 of Frontiers in Applied Mathematics, SIAM, Philadelphia, 1992.

[12] K. LOEWNER, *Über monotone Matrixfunktionen,* Math. Z. **38**, (1934) 177–216.

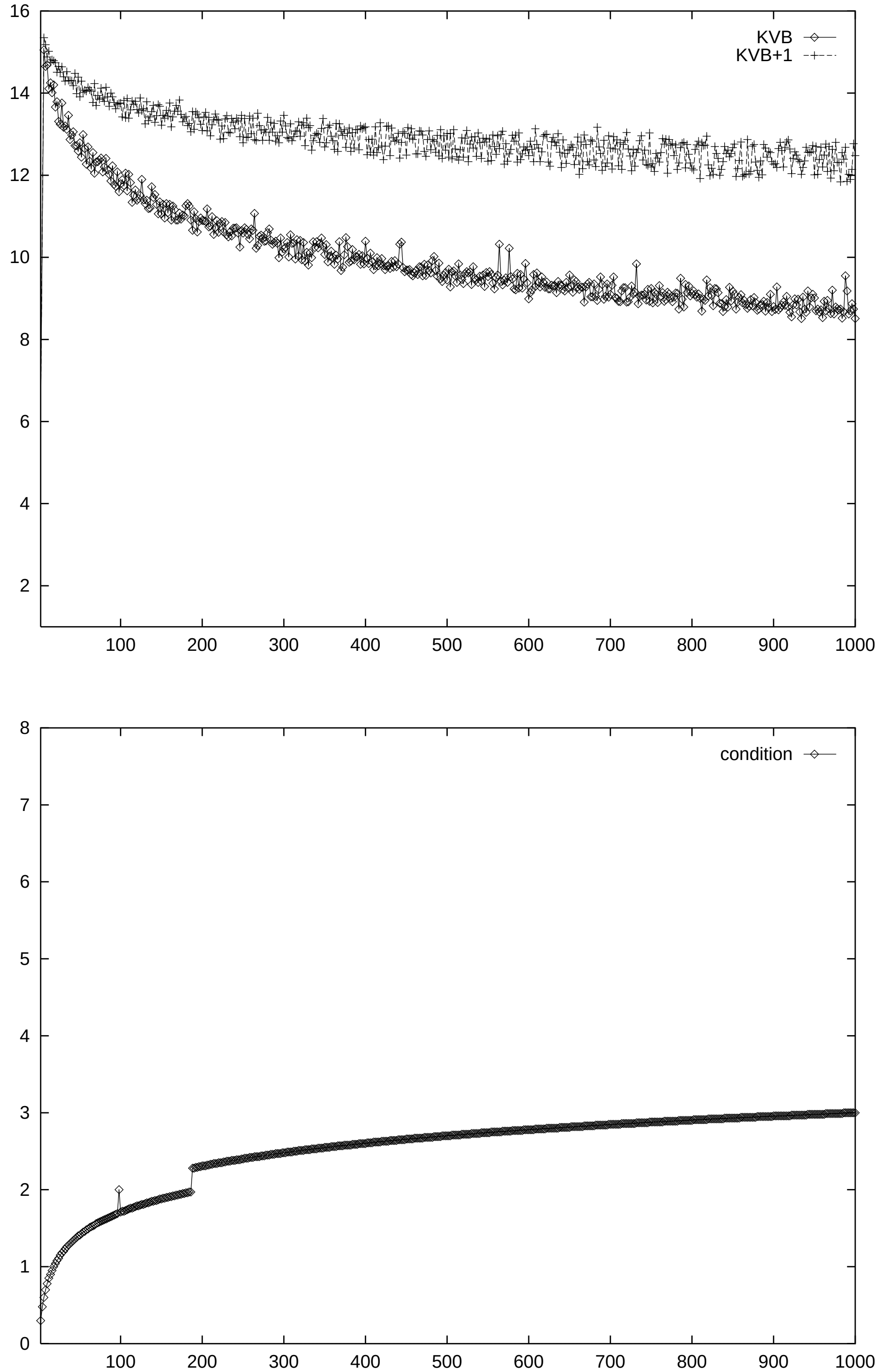

Figure 2: $-\log_{10} \dfrac{\|\hat{x}_{2n} - x_{2n}\|_\infty}{\|x_{2n}\|_\infty}$ and $\log_{10} \kappa_\infty(H_{2n})$ versus $2n$ (Example 2)

[13] M. VAN BAREL, A. BULTHEEL, *A look-ahead algorithm for the solution of block Toeplitz systems,* Accepted for publication in Linear Algebra Appl. (37 pages).

[14] M. VAN BAREL, A. BULTHEEL, *A general module theoretic framework for vector M-Padé and matrix rational interpolation,* Numer. Algorithms **3**, (1992) 451–462.

[15] M. VAN BAREL, Z. VAVŘÍN, *Inversion of a block Löwner matrix,* J. Comput. Appl. Math. **69**, (1996) 261–284.