ON THE 2-CONDITION NUMBER OF INFINITE HANKEL MATRICES OF FINITE RANK

by F.S.V. Bazán[1] and Ph.L. Toint[2]

Report 98/10 June 16, 1998

[1] Department of Mathematics,
Federal University of Santa Catarina,
Florianopolis, Santa Catarina 88040-900, Brasil.
Email: fermin@mtm.ufsc.br

[2] Department of Mathematics,
Facultés Universitaires ND de la Paix,
61, rue de Bruxelles, B-5000 Namur, Belgium.
Email: pht@math.fundp.ac.be

On the 2-Condition Number of Infinite Hankel Matrices of Finite Rank

F. S. V. Bazán * Ph. L. Toint

June 16, 1998

Abstract

Let H be an infinite Hankel matrix whose (i,j)-entry is h_{i+j-2}, where $\{h_k\}_{k=0}^{\infty}$ denotes a complex-valued sampled signal: $h_k = \sum_{l=1}^{n} r_l\, z_l^k,\ \ |z_l| < 1$. Then H has rank n and can be factorized as $H = WRW^T$, where W is an infinite Vandermonde matrix in $\mathbb{C}^{\infty\times n}$ with z_i^{j-1} as its (i,j)-entry and R a diagonal matrix containing the weights r_l. We derive upper bounds for the 2-condition number of H as functions of n, r_l and z_l, which show that this condition number is small whenever the z's are close to the unit circle but not extremely close to each other. An application in problems related to exponential modeling and system identification is presented and discussed.

Key words. Infinite Hankel matrices, singular values, eigenvalues, exponential modeling.

1 Introduction

Let H be an infinite Hankel matrix whose (i,j)-entry is h_{i+j-2}, that is

$$H = \begin{bmatrix} h_0 & h_1 & h_2 & \cdots \\ h_1 & h_2 & h_3 & \cdots \\ h_2 & h_3 & \cdots & \cdots \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix}, \tag{1.1}$$

where $\{h_k\}_{k=0}^{\infty}$ denotes a complex-valued sampled signal composed of n exponentials

$$h_k = \sum_{l=1}^{n} r_l\, z_l^k,\ \ k = 0, 1, \ldots, \tag{1.2}$$

where r_l, z_l are complex constants and the z's are known as *modes*. We assume $|z_l| < 1$ and $z_k \neq z_l$ for $k \neq l$. Then H has rank n and can be factorized as (see, for instance Gragg and Reichel, 1989)

$$H = W\,RW^T, \tag{1.3}$$

---

*Departamento de Matemática, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina 88040-900, Brasil (fermin@mtm.ufsc.br). This research was performed while the author was a visitor at FUNDP, Namur, Belgium, and was supported by CNPq, Brasil, grant 201299/96-8(NV).

where $R = \mathrm{diag}(r_1, r_2, \ldots, r_n)$ and W^T the infinite Vandermonde matrix

$$W^T = \begin{bmatrix} 1 & z_1 & z_1^2 & z_1^3 & \cdots \\ 1 & z_2 & z_2^2 & z_2^3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & z_n & z_n^2 & z_n^3 & \cdots \end{bmatrix}_{n\times\infty} . \tag{1.4}$$

Let ℓ^2 be the Hilbert space of infinite complex column vectors with finite 2-norm. Then H gives rise to a bounded linear operator of finite rank on ℓ^2 and thus, from linear operator theory, its Moore-Penrose pseudo inverse can be defined as the unique bounded operator $H^\dagger$ satisfying the conditions

$$HH^\dagger H = H, \;\; H^\dagger HH^\dagger = H^\dagger, \;\; (HH^\dagger)^* = HH^\dagger, \;\; (H^\dagger H)^* = H^\dagger H, \tag{1.5}$$

where the star symbol stands for the adjoint of the operator. The singular values of H are the square roots of the eigenvalues of H^*H and hence, since $\mathrm{rank}(H) = n$, its singular values which we denote by $\sigma_i(H)$, satisfy

$$\sigma_1(H) \geq \sigma_2(H) \geq \cdots \geq \sigma_n(H) > 0 = \sigma_{n+1}(H) = \sigma_{n+2}(H) = \cdots, \tag{1.6}$$

with $\sigma_1(H) = \|H\|_2$ and $\sigma_n(H) = \|H^\dagger\|_2^{-1}$, where $\|\cdot\|$ stands for the operator norm. Similarly, the infinite Vandermonde matrix W gives rise to another bounded linear operator of rank n, from $\mathbb{C}^n$ in ℓ^2, and its pseudo inverse $W^\dagger$ also can be defined via the above conditions. Taking this observation into account, we verify, using (1.5) for W, that

$$W^\dagger = (W^*W)^{-1}W^*. \tag{1.7}$$

Using this relation it is easy to see that ${W^T}^\dagger R^{-1} W^\dagger$ satisfies the conditions (1.5) and thus that

$$H^\dagger = {W^T}^\dagger R^{-1} W^\dagger. \tag{1.8}$$

All the above properties of pseudo-inverses are consequences of the well-developed Hilbert space theory for pseudo-inverses of bounded linear operators with closed range, which clearly also hold for matrices in $\mathbb{C}^{M\times N}$. The reader is referred to Ben-Israel and Greville (1974) for details about pseudo-inverses, and, in particular, to Theorems 2 and 3 of Chapter 8 therein.

The 2-condition number of H, $k_2(H)$, is defined as the ratio of the largest to the smallest non zero singular value of H, that is $k_2(H) = \sigma_1(H)/\sigma_n(H) = \|H\|_2\|H^\dagger\|_2$. This condition number is of interest in areas where the singular values of Hankel matrices of growing dimension are explicitly used, as in Hankel norm approximation problems in systems theory, (see, for instance, Glover, 1984), or in several signal processing applications such as exponential modeling or harmonic retrieval problems (see, for instance, Van Huffel, 1997). All these applications largely justify a study on upper bounds for $k_2(H)$, which is the main goal of this work.

Our analysis is based on the following modification of factorization (1.3):

$$H = UU^T, \;\; \text{where} \quad U = WR^{\frac{1}{2}}, \tag{1.9}$$

in which $R^{\frac{1}{2}}$ is an arbitrarily chosen square root of R. Notice that, since

$$U^{\dagger} = (U^{*}U)^{-1}U^{*} = (R^{\frac{1}{2}})^{-1}W^{\dagger}, \tag{1.10}$$

which can verified by checking (1.5) for U, then $H^{\dagger}$ can be rewritten as

$$H^{\dagger} = {U^{T}}^{\dagger}U^{\dagger}. \tag{1.11}$$

Now, observe that (1.9) and (1.11) imply that

$$k_2(H) \leq \|U\|_2^2\|U^{\dagger}\|_2^2 = [k_2(U)]^2, \tag{1.12}$$

and that the inequality in this bound holds as an equality when the modes z_l are real and the weights r_l are positive. This suggests that we first bound $k_2(U)$ in order to obtain upper bounds for $k_2(H)$. Thus, our strategy for deriving bounds for $k_2(H)$ consists of first deriving bounds for $k_2(U)$: we rely on clasical results about Jordan conditions numbers (see, for instance, Smith, 1967, or Wilkinson, 1965), which were used by Bazán (1998) for deriving an upper bound for $k_2(W)$. The bounds that we derive involve the parameters n, r_l and z_l and their quality essentially depends on the separation of the z's inside the unit circle. In particular, we show that whenever the z's are close to the unit circle but not extremely close to each other, the bounds for $k_2(H)$ are $\mathcal{O}(n^2)$. Observe that we could use the inequality

$$k_2(H) \leq k_2(R)k_2(W)^2$$

to perform our analysis, but the approach chosen here seem to give better results.

The paper is organized as follows. In Section 2 we consider finite sections of the Hankel matrix as Krylov matrices and some basic results arising from this identification are derived. Our upper bounds for $k_2(H)$ are presented in Section 3. Because the smallest non zero singular value of finite sections of H plays the role of a *threshold value* for separating signal from noise, in several signal processing applications (see, for instance De Moor, 1993), an application of our results illustrating that role is presented in Section 4. We analyze the choice of the dimension of the finite Hankel matrix that guarantees a satisfying separation of signal from noise. Section 5 finally presents some conclusions.

2 Viewing Hankel matrices as Krylov matrices and basic results

In what follows the singular values of a matrix A are denoted by $\sigma_i(A)$ and are arranged so that $\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_n(A)$. Also, as usual, $\|A\|_2$ and $\|A\|_F$ denote the spectral and the Frobenius norm of A, respectively. Leading $N \times N$ principal submatrices of the Hankel matrix H play a important role in our developments, and are denoted by H_N. A first consequence of the structure of these matrices is that if N is larger than or equal n, then H_N inherits its rank from H and also its Vandermonde decomposition (1.3), that is $\operatorname{rank}(H_N) = n$ and

$$H_N = W_N\, R\, W_N^T, \tag{2.1}$$

where W_N is the matrix consisting of the first N rows of W, and R is as above. It is well known that the entries of H satisfy a recurrence relation of order n of the form

$$h_k = f_{n-1}h_{k-1} + f_{n-2}h_{k-2} + \cdots + f_0 h_{k-n}, \quad k = n, n+1, \cdots \tag{2.2}$$

which generates the entire signal once the set of n initial values $\{h_0, h_1, \ldots, h_{n-1}\}$ are given (see Gantmacher (1959), vol. 2, p. 207). Furthermore, the modes z_l generating the entries in (1.2) are the roots of the polynomial

$$p_n(z) = z^n - f_{n-1}z^{n-1} - \cdots - f_1 z - f_0. \tag{2.3}$$

The coefficients f_i are uniquely determined from the recurrence relation and are referred to as *predictor parameters.* Recurrence relations of type (2.2) of order $N > n$ are also possible. In this case however, the predictor parameters are not uniquely determined, since they are computed from the rank-deficient underdetermined linear system

$$H_N f \overset{\text{def}}{=} [\mathbf{h}_0\ \mathbf{h}_1\ \cdots\ \mathbf{h}_{N-1}]f = \mathbf{h}_N, \tag{2.4}$$

where

$$\mathbf{h}_i \overset{\text{def}}{=} [h_i\ h_{i+1}\ \cdots\ h_{i+N-1}]^T, \quad i = 0, \ldots, N \quad \text{and} \quad f = [f_0\ f_1\ \cdots\ f_{N-1}]^T.$$

In spite of this, it is well known that the modes z_l still can be extracted from the roots of any polynomial $p_N(z)$ whose coefficients satisfy (2.4) (Wei and Majda, 1990). Let C denote the companion matrix corresponding to the polynomial

$$P_N(z) = z^N - f_{N-1}z^{N-1} - \cdots - f_1 z - f_0,$$

whose coefficients are the componentes of the minimum 2-norm solution of (2.4), i.e.

$$C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \\ f_0 & f_1 & \cdots & \cdots & f_{N-1} \end{bmatrix}. \tag{2.5}$$

The fact that the signal modes can be extracted from the zeros of $P_N(z)$ implies that there are n eigenvalues of C exactly coinciding with the n modes z_l. More precisely, it can be proved that

$$CW_N = W_N Z, \tag{2.6}$$

where $Z = \text{diag}(z_1, z_2, \ldots, z_n)$, which means that W_N is a matrix of right eigenvectors of C with associated eigenvalues z_l (see, for instance, Bazán and Bezerra, 1997). We may now rephrase the property that the coefficients f_i are sufficient to predict future values of the signal (see (2.2)) in terms of the matrix C: given two successive columns vectors of H_N, $\mathbf{h}_i$ and $\mathbf{h}_{i+1}$, then

$$\mathbf{h}_{i+1} = C\,\mathbf{h}_i, \; i \geq 0.$$

From this observation, it follows that the Hankel matrix H_N can be regarded as a Krylov matrix generated by C, that is

$$H_N = [\mathbf{h}_0 \; \mathbf{h}_1 \; \cdots \; \mathbf{h}_{N-1}] = [\mathbf{h}_0 \;\; C\mathbf{h}_0 \;\; \cdots \; C^{N-1}\mathbf{h}_0],$$

and that, since $\mathrm{rank}(H_N) = n$ for $N \geq n$, the associated Krylov subspace, which we denote by $\mathcal{H}_N$, is *invariant* under C. Let V be an $N \times n$ matrix with orthonormal columns spanning $\mathcal{H}_N$ and consider $C_{\mathcal{P}}$ to be the $n \times n$ matrix defined by

$$C_{\mathcal{P}} = V^* C V. \tag{2.7}$$

As our results crucially depend on the eigenvalue and singular value spectra of $C_{\mathcal{P}}$, we start by analyzing them. We first observe that $C_{\mathcal{P}}$ is the orthogonal projection of C onto $\mathcal{H}_N$ written in terms of the columns of V. Hence, since the columns of both V and W_N span $\mathcal{H}_N$, (2.6) ensures that Z is a eigenvalue matrix of $C_{\mathcal{P}}$ with $V^* W_N$ as a matrix of right eigenvectors, i.e.

$$C_{\mathcal{P}} = (V^* W_N) Z (V^* W_N)^{-1}. \tag{2.8}$$

In order to describe the singular value spectrum of $C_{\mathcal{P}}$ we shall analyse the eigenvalues of $C_{\mathcal{P}}^* C_{\mathcal{P}}$. Let $\mathcal{P}$ and I be the orthogonal projector onto $\mathcal{H}_N$ and the identity matrix of order n, respectively. Now, observe that, as $VV^* = W_N W_N^{\dagger} = \mathcal{P}$,

$$W_N^{\dagger} V V^* W_N = I = V^* W_N W_N^{\dagger} V,$$

and thus that

$$(V^* W_N)^{-1} = W_N^{\dagger} V. \tag{2.9}$$

(For details on pseudo-inverses and orthogonal projectors, see Björk, 1996.) Using the above relations involving $\mathcal{P}$ as well as (2.6), (2.8) and (2.9), we have that

$$C_{\mathcal{P}}^* C_{\mathcal{P}} = V^* W_N^{\dagger *} Z^* W_N^* V V^* W_N Z W_N^{\dagger} V = V^* C^* C V,$$

and hence, since

$$C^* C \;\; = \;\; [e_2 \; \ldots \; e_N \; f^{T*}] \begin{bmatrix} e_2^* \\ \vdots \\ e_N^* \\ f^T \end{bmatrix} = e_2 e_2^* + \cdots + e_N e_N^* + f^{T*} f^T,$$

it is easy to see that $C_{\mathcal{P}}^* C_{\mathcal{P}}$ can be expressed as a rank-two modification of the identity of the form

$$C_{\mathcal{P}}^* C_{\mathcal{P}} = I + xx^* - yy^*, \tag{2.10}$$

where we have introduced $x = V^* f^{T*}$ and $y = V^* e_1$. It then follows from this identity that

$$\mathrm{Tr}(C_{\mathcal{P}}^* C_{\mathcal{P}}) = \mathrm{Tr}(I + xx^* - yy^*) = n + \|x\|_2^2 - \|y\|_2^2, \tag{2.11}$$

and, since $\det(I + xx^*) = 1 + x^*x = 1 + \|x\|_2^2 \neq 0$, we also deduce that

$$\det(C_{\mathcal{P}}^* C_{\mathcal{P}}) = \det(I + xx^*)\det(I - (I + xx^*)^{-1} yy^*) = (1 + \|x\|_2^2)(1 - y^*(I + xx^*)^{-1} y).$$

After computing the inverse in this expression by applying the Sherman-Morrison formula, we obtain that

$$\det(C_{\mathcal{P}}^* C_{\mathcal{P}}) = (1 + \|x\|_2^2)(1 - \|y\|_2^2) + |x^*y|^2. \tag{2.12}$$

Now, since at least $n - 2$ eigenvalues of $C_{\mathcal{P}}^* C_{\mathcal{P}}$ are equal to one, which can be seen from (2.10), the remaining eigenvalues, λ_1 and λ_2, say, must satisfy

$$\lambda_1 + \lambda_2 + (n - 2) = \mathrm{Tr}(C_{\mathcal{P}}^* C_{\mathcal{P}}) \text{ and } \lambda_1 \lambda_2 = \det(C_{\mathcal{P}}^* C_{\mathcal{P}}). \tag{2.13}$$

These relations, (2.11) and (2.12) then imply that λ_1 and λ_2 must be roots of the polynomial

$$q(\lambda) = \lambda^2 - (2 + \|x\|_2^2 - \|y\|_2^2)\lambda + (1 + \|x\|_2^2)(1 - \|y\|_2^2) + |x^*y|^2, \tag{2.14}$$

which are given by

$$\lambda_1, \lambda_2 = \frac{2 + \|x\|_2^2 - \|y\|_2^2 \pm \sqrt{(\|x\|_2^2 + \|y\|_2^2)^2 - 4|x^*y|^2}}{2}. \tag{2.15}$$

We now notice that the smallest of these roots, λ_2, say, is smaller than one because the eigenvalues of $C_{\mathcal{P}}$ are smaller than one in modulus. We also observe that λ_1 cannot be smaller than one, otherwise we would have, from (2.15), that

$$(\|x\|_2^2 - \|y\|_2^2)^2 > (\|x\|_2^2 + \|y\|_2^2)^2 - 4|x^*y|^2.$$

This inequality holds only if $|x^*y| > \|x\|_2 \|y\|_2$, which is impossible as it contradicts the Cauchy-Schwarz inequality. On the other hand, since the vector of parameters f is the minimal norm solution of (2.4), it can be written as

$$f = H_N^\dagger \mathbf{h}_N = W_N^{T^\dagger} Z^N e, \tag{2.16}$$

where we have used (1.8) to compute $H_N^\dagger$ and $\mathbf{h}_N = W_N R Z^N e$, with e being the vector in $\mathbb{C}^n$ of all ones. Hence, one can see that $VV^* f^{T*} = f^{T*}$ and therefore that

$$\|x\|_2^2 = f^T VV^* f^{T*} = \|f\|_2^2, \;\; |x^*y| = |y^*x| = |e_1^* VV^* f^{T*}| = |f_0|,$$

where f_0 is the first component of f. Finally, since $\|V^* e_1\|_2 = \|VV^* e_1\|_2 = \|\mathcal{P} e_1\|_2$, we conclude that the singular values of $C_{\mathcal{P}}$ verify

$$\begin{aligned}
\sigma_1^2(C_{\mathcal{P}}) &= \frac{2 + \|f\|_2^2 - \|p_1\|_2^2 + \sqrt{(\|f\|_2^2 + \|p_1\|_2^2)^2 - 4|f_0|^2}}{2}, \\
\sigma_i^2(C_{\mathcal{P}}) &= 1, \;\; i = 2, n-1, \\
\sigma_n^2(C_{\mathcal{P}}) &= \frac{2 + \|f\|_2^2 - \|p_1\|_2^2 - \sqrt{(\|f\|_2^2 + \|p_1\|_2^2)^2 - 4|f_0|^2}}{2},
\end{aligned} \tag{2.17}$$

where p_1 is the first column of $\mathcal{P}$. All the above results are summarized in the following theorem.

Theorem 1 *Suppose that $C_{\mathcal{P}}$ is the matrix defined in (2.7) then it admits a spectral decomposition given by (2.8). Furthermore, its spectrum is formed by the signal modes z_l and its singular values $\sigma_i(C_{\mathcal{P}})$ satisfy (2.17).*

Similar results involving the singular values of companion matrices associated with *backward* predictor parameters have been given by Bazán and Toint (1998). The following theorem describes how the largest and the smallest of those singular values behave as the dimension of $C_{\mathcal{P}}$ tends to infinity.

Theorem 2 *Suppose $\sigma_1(C_{\mathcal{P}})$ and $\sigma_n(C_{\mathcal{P}})$ are described by (2.17). Then*

$$\lim_{N\to\infty} \sigma_1(C_{\mathcal{P}}) = 1, \quad \text{and} \quad \lim_{N\to\infty} \sigma_n(C_{\mathcal{P}}) = \prod_{i=1}^{n} |z_i|. \tag{2.18}$$

Proof: Since $\det(C_{\mathcal{P}}^* C_{\mathcal{P}}) = [\det(C_{\mathcal{P}})]^2 = \prod_{l=1}^{n} |z_l|^2$, the second part of (2.13) yields that

$$\sigma_1^2(C_{\mathcal{P}})\sigma_2^2(C_{\mathcal{P}}) = \prod_{l=1}^{n} |z_l|^2. \tag{2.19}$$

On the other hand, it is not difficult to see from (2.17) that

$$1 \le \sigma_1^2(C_{\mathcal{P}}) \le 1 + \|f\|_2^2. \tag{2.20}$$

We now analyze $\|f\|_2^2$ as $N \to \infty$. We first notice from (2.16) that

$$\|f\|_2 \le \|W_N^{\dagger}\|_2 \sqrt{n}\bar{\alpha}^N, \tag{2.21}$$

where $\bar{\alpha} = \max |z_l|$. Now, since

$$W_{N+1}^* W_{N+1} = W_N^* W_N + Z^{N*} e e^* Z^N,$$

observe that the singular values of W_{N+1} cannot be smaller than those of W_N, as $Z^{N*}ee^*Z^N$ is positive semidefinite. In other words, $\|W_N^{\dagger}\|$ decreases with N. Consequently, we deduce from (2.21) that $\|f\|_2 \to 0$ as $N \to \infty$, because $\bar{\alpha} < 1$. Using this result, the statements of the theorem immediately follow from (2.20) and (2.19).□

3 Upper bounds for $k_2(H)$

We now derive upper bounds for $k_2(H)$ by first deriving upper bounds for $k_2(U)$, where U is the infinite scaled Vandermonde matrix introduced in (1.12). In order to achieve our goal, we start with an analysis in the finite dimensional case and derive a bound for $k_2(U_N)$, where U_N is defined by $U_N = W_N R^{\frac{1}{2}}$.

Theorem 3 *Let U_N be the scaled $N \times n$ Vandermonde matrix and $z = [z_1 \ \dots \ z_n]^T$ and $r = [r_1 \ \dots \ r_n]^T$. Let p and q be indices between 1 and n, such that $|r_i z_i|$, $i = 1, \dots, n$, is maximum* *for* $i = p$ *and minimum for* $i = q$*. Then define* $\alpha = |z_p|,\ \ \rho = |r_p|,\ \ \beta = |z_q|,\ \ \gamma = |r_q|,$ *and* $k_R = \rho/\gamma$*. Also define,*

$$\delta = \min_{\substack{1 \leq i,j \leq n \\ i \neq j}} |z_i - z_j|, \tag{3.1}$$

$$\phi_N = \sqrt{\frac{1 + \alpha^2 + \cdots + \alpha^{2(N-1)}}{1 + \beta^2 + \cdots + \beta^{2(N-1)}}}, \quad \textit{and} \tag{3.2}$$

$$D_N^2 = (\sigma_1^2(C_{\mathcal{P}}) + \cdots + \sigma_n^2(C_{\mathcal{P}})) - (|z_1|^2 + \cdots + |z_n|^2), \tag{3.3}$$

where $\sigma_i(C_{\mathcal{P}})$ *are the singular values of* $C_{\mathcal{P}}$ *given by (2.17). Then, for all* $N \geq n$*, the spectral condition number of* U_N*,* $k_2(U_N) = \|U_N\|_2 \|U_N^\dagger\|_2$*, satisfies*

$$k_2(U_N) \leq \frac{1}{2}\left(\eta + \sqrt{\eta^2 - 4}\right),$$

where

$$\eta = \sqrt{k_R}\left[1 + \frac{D_N^2}{(n-1)\delta^2}\right]^{\frac{n-1}{2}} \frac{n}{2}\left(\phi_N + \frac{1}{k_R}\phi_N^{-1}\right) - n + 2. \tag{3.4}$$

Proof: The proof is based on the crucial observation that the conditioning of the eigenvalue problem associated with the matrix $C_{\mathcal{P}}$ (see Theorem 1) is ultimately governed by the conditioning of U_N. Let $u_i = V^* U_N^{\dagger *} e_i$ and $v_i = V^* U_N e_i$ where e_i is the i-th canonical basis vector in $\mathbb{C}^n$. It follows that they are left and right eigenvectors of $C_{\mathcal{P}}$ associated with the eigenvalue z_i, respectively, because of (2.8) and (2.9), and that they satisfy $u_i^* v_i = 1$, by construction. Furthermore, we have that $\|v_i\|_2^2 = \|V v_i\|_2^2 = \|V V^* U_N e_i\|_2^2 = \|U_N e_i\|_2^2$, $\|u_i\|_2^2 = \|V u_i\|_2^2 = \|V V^* U_N^{\dagger *} e_i\|_2^2 = \|U_N^{\dagger *} e_i\|_2^2$, because VV^* is the orthogonal projector onto $\mathcal{H}_N$ and the columns of U_N span this subspace. This yields that

$$\|U_N\|_F^2 = \sum_{i=1}^n \|v_i\|_2^2 \quad \text{and} \quad \|U_N^\dagger\|_F^2 = \sum_{i=1}^n \|u_i\|_2^2.$$

Using these observations, we obtain from Theorem 5 by Smith, 1967, that

$$\|u_i\|_2 = \frac{1}{|s_i|}\frac{1}{\|v_i\|_2} \leq \left[1 + \frac{D_N^2}{(n-1)\delta_i^2}\right]^{\frac{n-1}{2}} \frac{1}{\|v_i\|_2}, \quad i = 1, 2, \ldots, n,$$

with

$$\delta_i = \min_{\substack{1 \leq j \leq n \\ i \neq j}} |z_i - z_j|,$$

and where

$$s_i = \frac{u_i^* v_i}{\|u_i\| \|v_i\|}, \tag{3.5}$$

is the condition number of the i-th eigenvalue (see Wilkinson, 1965, page 69). From this, it follows that

$$\|U_N^\dagger\|_F^2 \leq \sum_{i=1}^n \left[1 + \frac{D_N^2}{(n-1)\delta_i^2}\right]^{n-1} \frac{1}{\|v_i\|_2^2} \leq \left[1 + \frac{D_N^2}{(n-1)\delta^2}\right]^{n-1} \sum_{i=1}^n \frac{1}{\|v_i\|_2^2}, \tag{3.6}$$

where δ is defined by (3.1). This implies that

$$\|U_N^\dagger\|_F^2 \|U_N\|_F^2 \leq \left[1 + \frac{D_N^2}{(n-1)\delta^2}\right]^{n-1} \sum_{i=1}^n \|v_i\|_2^2 \sum_{i=1}^n \frac{1}{\|v_i\|_2^2}. \tag{3.7}$$

Now observe that, if we define $Q = \text{diag}(\|v_1\|_2^2, \|v_2\|_2^2, \ldots, \|v_n\|^2)$, the product of the two sums in the right-hand-side of this inequality may be rewriten as

$$\sum_{i=1}^n \|v_i\|_2^2 \sum_{i=1}^n \frac{1}{\|v_i\|_2^2} = e^* Q e \, e^* Q^{-1} e,$$

and that this quantity may then bounded from above by using Kantorovic's inequality (Horn and Johnson, 1990, pag. 444), which implies that

$$e^* Q e \, e^* Q^{-1} e \leq \left[\frac{n}{2}\left(\frac{v_{max}}{v_{min}} + \frac{v_{min}}{v_{max}}\right)\right]^2,$$

where $v_{max} = \max \|v_i\|_2$ and $v_{min} = \min \|v_i\|_2$. Using the previous inequality and the notations introduced in the theorem, and since

$$\|v_i\|_2^2 = \|W_N R^{\frac{1}{2}} e_i\|_2^2 = |r_i|(1 + |z_i|^2 + \cdots + |z_i|^{2(N-1)}),$$

(3.7) gives that

$$k_F(U_N) \leq \sqrt{k_R}\left[1 + \frac{D_N^2}{(n-1)\delta^2}\right]^{\frac{n-1}{2}} \frac{n}{2}(\phi_N + \frac{1}{k_R}\phi_N^{-1}). \tag{3.8}$$

We now recall a property regarding Jordan condition numbers (see Smith, 1967, Theorem 1) which states that, if $A = X\Lambda X^{-1}$ is a spectral decomposition of $A \in \mathbb{C}^{n\times n}$ and all eigenvalues of A are simple, then

$$n - 2 + k_2(X) + [k_2(X)]^{-1} \leq k_F(X). \tag{3.9}$$

Solving this inequality for $k_2(X)$ and adapting the result to our context, we obtain

$$k_2(U_N) \leq \frac{1}{2}\left[k_F(U_N) - n + 2 + \sqrt{(k_F(U_N) - n + 2)^2 - 4}\right],$$

and the proof concludes by substituting (3.8) in this inequality. □

Since we have that $H_N = U_N U_N^T$ by (2.1), we have then that $k_2(H_N) \leq [k_2(U_N)]^2$. Hence, using the results of the above theorem we obtain the following consequence.

Corollary 4 *Suppose $N \geq n$. Then*

$$k_2(H_N) \leq \frac{1}{4}\left(\eta + \sqrt{\eta^2 - 4}\right)^2, \tag{3.10}$$

where η is as above.

We immediately notice that, if $R = I$ (i.e, all weights r_l are one), then we obtain an upper bound for the 2-condition number of the unscaled Vandermonde matrix W_N, which improves that of Bazán (1998). In spite of the obtained improvement, both bounds strongly depend on D_N^2 and

δ in that small values of the bounds are ensured only when $D_N^2 \le (n-1)\delta^2$. While δ^2 measures the separation of the modes z_l inside the unit circle, D_N^2, known as the *departure from normality*, measures how near is $C_{\mathcal{P}}$ from being a normal matrix (see, for instance, Golub and Van Loan (1996) pag. 314). Numerical examples showing the dependence of the bound on those quantities are presented and discussed in the next section. However, before presenting them, we pursue our task to derive upper bounds for $k_2(H)$. We first state an auxiliary result.

Lemma 5 *Let U_N be the $N \times n$ scaled Vandermonde matrix as above. Define two sequences of real numbers $\{a_N\}$, $\{b_N\}$, $N \ge n$, by $a_N = \|U_N^\dagger\|_2$ and $b_N = k_2(U_N) = \|U_N\|_2\|U_N^\dagger\|_2$. Then,*

- (a) *a_N decreases monotonically with N;*
- (b) $\lim_{N\to\infty} b_N = k_2(U)$.

Proof: The proof of part (a) is similar to that presented in Section 2 where we proved that $\|W_N^\dagger\|_2$ decreases with N (see Theorem 2). To prove (b), let $\{\overline{U}_N\}$ be the sequence of infinite matrices with zeros everywhere except in the first N rows whose entries coincide with those of U_N, i.e,

$$\overline{U}_N = \begin{bmatrix} U_N \\ 0 \\ \vdots \end{bmatrix}.$$

Obviously, $\|\overline{U}_N\|_2 = \|U_N\|_2$ and $\|\overline{U}_N^\dagger\|_2 = \|U_N^\dagger\|_2$. Next, partition U as

$$U = \begin{bmatrix} U_N \\ C \end{bmatrix} = \overline{U}_N + \begin{bmatrix} 0 \\ C \end{bmatrix}, \tag{3.11}$$

where $C = [Z^N e \;\; Z^{N+1}e \;\; \cdots]^T R^{\frac{1}{2}}$, and notice that the first part of (1.10) can be rewritten as

$$U_N^* U_N U^\dagger + C^* C U^\dagger = [U_N^* \;\; 0 \cdots] + [0 \;\; C^*].$$

This can be rewritten again as

$$U^\dagger + (U_N U_N^*)^{-1}(C^*C)U^\dagger = \overline{U}_N^\dagger + (U_N U_N^*)^{-1}\,[0 \;\; C^*],$$

which, using the fact that $U_N^\dagger U_N^{\dagger *} = (U_N^* U_N)^{-1}$, yields that

$$U^\dagger - \overline{U}_N^\dagger = -U_N^\dagger U_N^{\dagger *}(C^*C)U^\dagger + U_N^\dagger U_N^{\dagger *}\,[0 \;\; C^*].$$

Hence, taking into account that both $\|U^\dagger\|_2$ and $\|U_N^\dagger\|_2$ are bounded ($U^\dagger$ is of finite rank and $\|U_N^\dagger\|_2$ decreases by part (a)), that $\|C^*C\|_2 = \|C\|_2^2$ and that

$$\|C\|_2 \le \sqrt{n}\,\sqrt{\rho}\,\alpha^N \|U\|_2 \to 0$$

as $N \to \infty$, because $\alpha < 1$, we deduce that

$$\|U^\dagger - \overline{U}_N^\dagger\|_2 \le \|U^\dagger\|_2\|C\|_2^2\|U_N^\dagger\|_2^2 + \|C\|_2\|U_N^\dagger\|_2^2 \;\to\; 0$$

whenever $N \to \infty$, implying that $\overline{U}_N^{\dagger} \to U^{\dagger}$. Hence $\|\overline{U}_N^{\dagger}\|_2$ converges to $\|U^{\dagger}\|_2$. Furthermore, (3.11) implies that

$$\|U - \overline{U}_N\| = \|C\| \to 0,$$

which then yields that $\|\overline{U}_N\|_2 \to \|U\|_2$. This implies that $b_N = \|U_N\|_2\|U_N^{\dagger}\|_2 = \|\overline{U}_N\|_2\|\overline{U}_N^{\dagger}\|_2$ converges to $k_2(U)$, as requested. □

Theorem 6 *Let H be the infinite Hankel matrix introduced in (1.1). Let α, β, ρ, γ, k_R and δ be as in Theorem 3. Then the 2-condition number of H, $k_2(H) = \|H\|_2\|H^{\dagger}\|_2$, satisfies*

$$k(H) \leq \frac{1}{4}\left(\overline{\eta} + \sqrt{\overline{\eta}^2 - 4}\right)^2, \tag{3.12}$$

where

$$\overline{\eta} = \sqrt{k_R}\left[1 + \frac{n - 1 + \prod_{i=1}^{n}|z_i|^2 - \sum_{i=1}^{n}|z_i|^2}{(n-1)\delta^2}\right]^{\frac{n-1}{2}} \frac{n}{2}\left[\sqrt{\frac{1-\beta^2}{1-\alpha^2}} + \frac{1}{k_R}\sqrt{\frac{1-\alpha^2}{1-\beta^2}}\right] - n + 2. \tag{3.13}$$

Proof: Since by (1.12) and Lemma 5,

$$k_2(H) \leq [k_2(U)]^2 = \lim_{N\to\infty}[k_2(U_N)]^2, \tag{3.14}$$

it is sufficient to take limit in (3.4) when N tends to infinity. In order to compute this limit, we observe that

$$\lim_{N\to\infty} D_N^2 = n - 1 + \prod_{i=1}^{n}|z_i|^2 - \sum_{i=1}^{n}|z_i|^2, \quad \text{and} \quad \lim_{N\to\infty}\phi_N = \sqrt{\frac{1-\beta^2}{1-\alpha^2}}, \tag{3.15}$$

the first because of (3.3) and (2.18), and the second because of (3.2) and the fact that α and β are both smaller than one. The desired result then obtained by substituting these values in (3.14). □

This shows that the quality of the upper bound depends on the closenees of $|z_l|$ to 1 and on the separation of the modes themselves inside the unit circle. The following corollary gives more information on the dependence of the bounds on the distribution of the modes inside the unit circle.

Corollary 7 *Assume that the modes z_l satisfy $1 - \bar{\beta}^2 \leq \delta^2$ and $\bar{\alpha}^n \leq \bar{\beta}$, where $\bar{\alpha} = \max|z_l|$ and $\bar{\beta} = \min|z_l|$. Then*

$$k_2(H) \leq \left[\sqrt{k_R}\,2^{\frac{n-3}{2}}n\left(\sqrt{\frac{1-\beta^2}{1-\alpha^2}} + \frac{1}{k_R}\sqrt{\frac{1-\alpha^2}{1-\beta^2}}\right) - n + 2\right]^2. \tag{3.16}$$

Proof: First, notice from (3.12) that

$$k_2(H) \leq \overline{\eta}^2. \tag{3.17}$$

Now, since $\bar{\alpha}$ and $\bar{\beta}$ are both smaller than one, then

$$n - 1 + \prod_{i=1}^{n}|z_i|^2 - \sum_{i=1}^{n}|z_i|^2 \leq (n-1) + \bar{\alpha}^{2n} - n\bar{\beta}^2 = (n-1)(1-\bar{\beta}^2) + \bar{\alpha}^{2n} - \bar{\beta}^2 \leq (n-1)(1-\bar{\beta}^2),$$

since by assumption $\bar{\alpha}^n \leq \bar{\beta}$. But, as we have assumed $1 - \bar{\beta}^2 \leq \delta^2$, the proof of the corollary follows from (3.17) after using this inequality into the definition of $\overline{\eta}$ in (3.13). □

Consequently, if the above conditions are satisfied, one obtains a bound which is $\mathcal{O}(n^2)$, predicting that the conditioning itself may be reasonable provided n is not so large. Theorem 6 indicates that the situation is even more favourable when δ^2 is not very small and $\bar{\alpha} \approx \bar{\beta} \approx 1$. The assumption $\bar{\alpha} \approx \bar{\beta} \approx 1$ is not only mathematically convenient, but as we shall now illustrate, occurs in signal processing applications featuring slightly damped signals.

4 A Signal Processing Application

In several signal processing applications one is interested in retrieving parameters such as frequencies, plane waves, dampings, etc, from a finite set of perturbed data. This data arises as $\tilde{h}_k = h_k + \epsilon_k$, $k = 0, 1, \ldots, L-1$, where h_k is an unknown sampled signal assumed to be of the form

$$h_k = \sum_{l=1}^{n} r_l \, e^{(d_l + i\omega_l)k\Delta t}, \tag{4.1}$$

with $r_l \in \mathbb{C}^n$, $i = \sqrt{-1}$, $d_l < 0$, $\omega_l \in \mathsf{R}$, Δt being the sampling interval, n the number of modes $z_l^k = e^{(d_l + i\omega_l)k\Delta t}$ contained in the signal and where ϵ_k is the noise. The aim is then to detect the number of spectral components of the signal as well as to compute, as accurately as possible, estimates of r_l, α_l and ω_l, even if the data is relatively noisy. We refer the reader to Van Huffel (1997) for a variety of applications and methods in which this problem is relevant.

From here on, leading $M \times N$ submatrices of H are denoted by $H_{M\times N}$. Two well-known approaches are often used to solve the problem: methods that extract the parameters from the roots of *large* polynomials, as described in Section 2, and methods based on estimates of the so-called *signal spaces* (the column or row space of $H_{M\times N}$, De Moor, 1993). Crucial for these approaches is the detection of n (the rank of H) and the estimation of the chosen signal space, both informations being extracted from a Hankel matrix which is generally of full rank: $\tilde{H}_{M\times N} = H_{M\times N} + E$, where E contains the noise. In practice, this is accomplished by a heuristic criterion for looking for a break in the pattern of singular values of $\tilde{H}_{M\times N}$, attributing the larger ones to the signal and the smaller ones to the noise. Similarly, the subspace spanned by the singular vectors associated with the set of large singular values is used as an estimate of the signal subspace. For a discussion concerning conditions on the noise matrix E that allow to recover the row signal space from the SVD of $\tilde{H}_{M\times N}$, see De Moor (1993).

We are now interested in the question to determine the best choice of M and N for the purpose of separating the n signal singular values from those associated with the noise, in the situation where the data is arranged in a Hankel matrix $\tilde{H}_{M\times N}$ such that $M + N = L + 1$, where L is fixed, $M, N \geq n$, and the noise matrix E is treated simply algebrically, i.e., with no assumption on the nature of the noise. As singular value theory implies that this separation is best carried out when $\|E\|_2 \ll \sigma_n(H_{M\times N})$ (which ensures a well visible gap between $\sigma_n(\tilde{H}_{M\times N})$ and $\sigma_{n+1}(\tilde{H}_{M\times N})$), we could focus on an analysis of the pair (M, N) that maximizes $\sigma_n(H_{M\times N})$ as a function the dimensions. This maximization is also of interest because the quality of the approximate signal subspace estimated from noisy measurements typically depends on expressions of the type $\|E\|_2/\sigma_n(H_{M\times N})$ (see, for instance, Theorem 2.1 in Fierro, Vanhamme and Van Huffel, 1997). However, this analysis is difficult and we conduct it on a model of $\sigma_n(H_{M\times N})$ instead. To construct this model we start by writing $H_{M\times N} = W_M R W_N^T$. Hence, using (1.8) we derive that $H^\dagger_{M\times N} = {W_N^T}^\dagger R^\dagger W_M^\dagger$, and therefore that

$$\sigma_n(H_{M\times N}) = {\|H^\dagger_{M\times N}\|_2}^{-1} \geq \|W_N^\dagger\|_2^{-1}\|R^\dagger\|_2^{-1}\|W_M^\dagger\|_2^{-1} = \bar{\gamma}\ \sigma_n(W_M)\sigma_n(W_N),$$

with $\bar{\gamma} = \min|z_l|$. Imposing the constraint that $M + N = L + 1$, this inequality becomes

$$\sigma_n(H_{M\times N}) \geq \bar{\gamma}\sigma_n(W_M)\sigma_n(W_{L+1-M}) \ \text{ for } \ n \leq M \leq L+1-n. \tag{4.2}$$

We then choose to model the left-hand side of this inequality by an estimate of its lower bound on the right-hand side. Observe next that for $M \approx L+1-n$ (which implies $N \approx n$), this bound is small, since $\sigma_n(W_{L+1-M}) \approx 0$ as W_{L+1-M} is almost a square Vandermonde matrix, which is generally ill-conditioned. This is in contrast with the fact that $\sigma_n(W_M)$ increases with M (ensured by Lemma 5). This balancing effect between M and N suggests that the Hankel matrix should not be chosen too overdetermined. By symmetry, the same reasoning applies in the case $M \approx n$ (which implies $N \approx L+1-n$), i.e, $H_{M\times N}$ should not be chosen too underdetermined. Hence, if the aim is to maximize the bound (4.2), then the Hankel matrix should not be chosen *too rectangular*. Indeed, we shall indicate that, under reasonable conditions, a satisfactory choice is $M \approx N$. Let $\xi(t)$ be an estimate of $\sigma_n(W_N)$ and consider a model of the form

$$\hat{\sigma}_n(H)(t) = \xi(t)\xi(L+1-n-t), \quad n \leq t \leq L+1-n, \tag{4.3}$$

where we impose $\xi(t)$ to be a real function defined in the interval $[n, L+1-n]$, satisfying

(i) $\xi(t) > 0$ for all $t \in [n, L+1-n]$

(ii) $\xi(t)$ is a non decreasing function of t;

(iii) $\xi(t)$ is twice differentiable and $\xi''(t) \leq 0\ \ \forall t$ in some neighbourhood of $t = \frac{1}{2}(L+1)$.

Observe as (i)–(ii) mimics the behaviour of $\sigma_n(W_N)$, as desired. Condition (iii) requires that $\xi(t)$ is locally concave around $t = \frac{1}{2}(L+1)$, which seems to be a property commonly obtained for $\sigma_n(W_N)$ in numerical experiments, provided sufficient data points are used.

Taking derivatives on both sides of (4.3), we obtain that

$$\hat{\sigma}'_n(H)(t) = -\xi(t)\xi'(L+1-t) + \xi'(t)\xi(L+1-t),$$

and hence that, $\hat{\sigma}'_n(H)(\frac{1}{2}(L+1)) = 0$, i.e, $t = \frac{1}{2}(L+1)$ is a first-order critical point of $\hat{\sigma}_n(H)(t)$. We now prove that it reaches a maximum. In fact, taking second derivatives, we have that

$$\hat{\sigma}''_n(H)(t) = \xi(t)\xi''(L+1-t) - 2\xi'(t)\xi'(L+1-t) + \xi''(t)\xi(L+1-t).$$

But, since $\xi(\frac{1}{2}(L+1)) > 0$, $\xi'(\frac{1}{2}(L+1)) \geq 0$ and $\xi''(\frac{1}{2}(L+1)) \leq 0$, by assumptions $(i) - (iii)$, we conclude that $\hat{\sigma}''_n(H)(\frac{1}{2}(L+1)) \leq 0$, ensuring that $\sigma_n(H)(t)$ reaches a maximum at $t = \frac{1}{2}(L+1)$.

In order to translate this result in terms of the singular value itself, we make the simplifying assumption that

$$\hat{\sigma}_n(H)(t)_{|_{t=j}} = \sigma_n(H_{j\times(L+1-j)}), \quad n \le j \le L+1-n. \tag{4.4}$$

Then the following result holds.

Theorem 8 *Let* $H_{M\times N}$ *be the Hankel matrix as above where* $M+N=L+1$, *and* L *is given. Then, provided* L *is odd and (4.4) holds, we have that* $\sigma_n(H_{M\times N})$ *is maximized when* $M=N=\frac{1}{2}(L+1)$ *and the following lower bound holds*

$$\sigma_n(H_{M\times M}) \ge K\left[1+\frac{D_M^2}{(n-1)\delta^2}\right]^{1-n}(1-e^{2d\Delta tM}),$$

where $d=\max|d_l|$ *and* $K=\dfrac{1}{\bar{\gamma}n(1-e^{2d\Delta t})}$ *and* $\bar{\gamma}$ *as above.*

Proof: Only the last part remains to be proved. Using the fact that $\|W_M^\dagger\|_2 \le \|W_M^\dagger\|_F$, from (3.6) we obtain that

$$\|W_M^\dagger\|_2^2 \le \left[1+\frac{D_M^2}{(n-1)\delta^2}\right]^{n-1}\frac{n}{1+e^{2d\Delta t}+\cdots+e^{2d\Delta t(M-1)}},$$

where $d=\max|d_l|$. The conclusion follows after inserting this inequality into (4.2), where we set $N=M=L+1-M$. $\square$

This theorem shows that if the continuous model we have just analyzed agrees with $\sigma_n(H_{M\times N})$ as described in (4.4), then the n-th singular value is maximized when $M=N$. Hence, if we assume that the noise is not high enough to dominate the signal, then the best gap between $\sigma_n(\tilde{H}_{M\times N})$ and $\sigma_n(\tilde{H}_{M\times N})$, say, $g_n(M,N)=\sigma_n(\tilde{H}_{M\times N})-\sigma_{n+1}(\tilde{H}_{M\times N})$, is likely to happen for $M=N$, and thus, choosing square Hankel matrices seems reasonable.

However, we are aware that, even if numerical experiments very often show that $\sigma_n(H_{M\times N})$ is maximized when we use square Hankel matrices, a model of the form (4.3) that exactly satisfies the assumption (4.4) is very difficult to construct. If we allow this latter requirement to hold only approximately, we may expect that the choice $M\approx N$ remains reasonable. On the other hand, because we have no control over $\sigma_n(\tilde{H}_{M\times N})$ and $\sigma_{n+1}(\tilde{H}_{M\times N})$, the choice $M=N$ may not be optimal in all cases, which also supports the relaxed rule $M\approx N$.

In order to illustrate the above observations as well as the behaviour of the bounds for $k_2(H)$, we have carried out a number of numerical experiments of which the most relevant are presented below.

4.1 Example 1: A simulated mechanical system

In this example we consider a simulated mechanical system with four degrees of freedom, whose impulse response function is defined by

$$h(t)=0.2e^{-0.06t}\sin(25t)+0.16e^{-0.056t}\sin(27t)+0.12e^{-0.09t}\sin(18t)+0.15e^{-0.2t}\sin(15t).$$

As this is a real signal, its sampled version (4.1) is composed of 8 exponentials, which implies that the associated Hankel matrix H has rank 8 (i.e. $n = 8$). We first illustrate the role that $\sigma_8(H_{M\times N})$ plays in the crucial problem of choosing a pair (M, N) that maximizes the gap $g_8(M, N)$. To this end, we analyze a set of 255 noise corrupted samples of the signal (i.e. $L = 255$), where we use zero-mean gaussian noise and $\Delta t = 0.05$. We next compute $\sigma_8(H_{M\times N})$, $\sigma_8(\tilde{H}_{M\times N})$, $\sigma_9(\tilde{H}_{M\times N})$ and $\|E\|$ for all pairs (M, N) such that $M + N = 256$. Results of this experiment for a noise level $\|\mathbf{e}\|_2/\|\mathbf{h}\|_2 \approx 40\%$ (corresponding to a standard deviation of 0.05), where $\mathbf{h}$ and $\mathbf{e}$ are vectors respectively containing the pure signal and noise, are shown in Figure 1-(b). Notice in this figure that the choice $M = N$ yields the best gap $g_8(M, N)$. This not only agrees with our theory but also emphasizes the importance of choosing the dimensions well: if we chose $M = 30$, no clear gap appears. Figure 1-(a) illustrates the behaviour of $\sigma_8(H_{M\times N})$ and a discrete approximate model of the form $K_1\sigma_8(W_M)\sigma_8(W_{L+1-M})$ where K_1 is a suitably chosen positive constant. Notice that the maximum value for both the true and modeled n-th singular value occur for $M = N = 128$.