SIAM J. MATRIX ANAL. APPL.
Vol. 21, No. 2, pp. 679–693

CONDITIONING OF RECTANGULAR VANDERMONDE MATRICES WITH NODES IN THE UNIT DISK*

FERMÍN S. V. BAZÁN†

Abstract. Let $W_N = W_N(z_1, z_2, \ldots, z_n)$ be a rectangular Vandermonde matrix of order $n \times N$, $N \geq n$, with distinct nodes z_j in the unit disk and z_j^{k-1} as its (j,k) entry. Matrices of this type often arise in frequency estimation and system identification problems. In this paper, the conditioning of W_N is analyzed and bounds for the spectral condition number $\kappa_2(W_N)$ are derived. The bounds depend on n, N, and the separation of the nodes. By analyzing the behavior of the bounds as functions of N, we conclude that these matrices may become well conditioned, provided the nodes are close to the unit circle but not extremely close to each other and provided the number of columns of W_N is large enough. The asymptotic behavior of both the conditioning itself and the bounds is analyzed and the theoretical results arising from this analysis verified by numerical examples.

Key words. Vandermonde matrices, singular values, almost normal matrices, exponential modeling

AMS subject classifications. 15A12, 65F35, 93A30

PII. S0895479898336021

1. Introduction.

Let $W_N = W_N(z_1, z_2, \ldots, z_n)$ be a Vandermonde matrix of order $n \times N$, $N \geq n$, with distinct nonzero complex nodes in the unit disk and z_j^{k-1} as its (j,k) entry, that is,

$$W_N = W_N(z_1, z_2, \ldots, z_n) = \begin{bmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^{N-1} \\ 1 & z_2 & z_2^2 & \cdots & z_2^{N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & z_n & z_n^2 & \cdots & z_n^{N-1} \end{bmatrix}, \tag{1.1}$$

and let $\kappa_2(W_N) = \|W_N\|\|W_N^\dagger\|$ be its 2-norm condition number, where † stands for Moore–Penrose pseudoinverse. These matrices arise in approximation theory and polynomial interpolation, but their most frequent use is in the framework of signal processing.

The purpose of this work is to derive bounds for $\kappa_2(W_N)$ and to analyze their behavior for fixed $n, z_1, \ldots, z_n$ and increasing N. Motivation for this analysis can be encountered in problems related to frequency estimation, parameter identification, and quantitative analysis of time domain nuclear magnetic resonance (NMR) data, among others; see, for instance, Van Huffel [17], Van Huffel et al. [18], and Bazán and Bavastri [1]. For illustration purposes, suppose we have a set of data h_l ($l = 0, 1, \ldots M-1$) involving $2n$ unknown parameters $r_j, z_j \in \mathbb{C}$ through the model

$$h_l = r_1 z_1^l + r_2 z_2^l + \cdots + r_n z_n^l,$$

*Received by the editors March 20, 1998; accepted for publication (in revised form) by N. J. Higham April 15, 1999; published electronically January 25, 2000. This research was performed while the author was a visitor at FUNDP, Namur, Belgium, and was supported by CNPq, Brazil, by grant 201299/96-8(NV).

http://www.siam.org/journals/simax/21-2/33602.html

†Departamento de Matemática, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina 88040-900, Brazil (fermin@mtm.ufsc.br).

and suppose we wish to determine these parameters from the available data. This is a separable problem that can be solved as follows. First, note that if all z_j are available then the r_j can be estimated by solving the least squares problem

$$\min \|W_M^T r - h\|, \tag{1.2}$$

where T denotes the transpose of a matrix, W_M is as before, and $h \in \mathbb{C}^M$ contains the data h_l. The sensitivity of the r_j to perturbations in the data is governed by $\kappa_2(W_M)$. If both r_j and z_j are unknown instead, then one starts by finding first the z_j. For instance, these z_j can be found from the roots of a polynomial $p_N(z) = a_0 + a_1 z + \cdots + a_{N-1}z^{N-1} + z^N$, whose coefficients are computed by solving a least squares problem of type

$$\min \|Ha + b\|, \tag{1.3}$$

where H is an $L \times N$ Hankel matrix with h_{i+j-2} as its (i,j) entry, $n \leq L, N \leq M-n$, and $b = [h_N, h_{N+1}, \ldots, h_{N+L-1}]^T$ (Wei and Majda [19]). Once the z_j are found, the r_j are readily computed by solving (1.2). In this case, one can prove that the sensitivity of the roots z_j to perturbations in the coefficients depends on $\|W_N^\dagger\|$ (see, for instance, Bazán, Toint, and Zambaldi [4]), and once more $\kappa_2(W_N)$ is important.

Although there are many other applications in signal processing where $\kappa_2(W_N)$ plays a crucial role, little is known about the behavior of $\kappa_2(W_N)$ for $N > n$, which is in contrast with the well-known fact that square Vandermonde matrices are in general ill conditioned (see, e.g., Gautschi [7], Gautschi and Inglese [8], Cordoba, Gautschi, and Ruscheweyh [5], and Tyrtyshnikov [15]). This justifies our interest to analyze this condition number.

Rectangular Vandermonde matrices also appear in trigonometric polynomial interpolation problems, but we shall restrict ourselves to analyzing the conditioning of those Vandermonde matrices that appear implicitly in frequency estimation and parameter identification. The term *implicitly* is used because in these applications the nodes z_j are known only approximately: the z_j appear after preliminary computations involving data h_l that is noise corrupted. Thus, what we pursue is a theoretical prediction of the sensitivity of these problems to perturbations in the data rather than an effective computation of $\kappa_2(W_N)$.

An interesting feature often numerically observed in the above problems is that $\kappa_2(W_N)$ is substantially better than the bad condition number $\kappa_2(W_n)$ when N is large enough, but a mathematical explanation for this phenomenon is still lacking. In this paper, we intend to substantiate this observation by exhibiting lower and upper bounds for $\kappa_2(W_N)$ which say much about the condition number itself. The bounds depend on n, z_j, and N, and their quality essentially depends on the separation of the nodes in the unit disk. In particular, we show that the numbers $\kappa_2(W_N)$ become rather small for large N, provided the nodes are not extremely close to each other and of modulus not much smaller than one. On the other hand, to evaluate the quality of the bounds as N grows, we analyze the asymptotic value of both the bounds and the condition number itself as N grows infinitely. As a by-product, we obtain that if all $|z_j| = 1$, then the asymptotic value of the bounds is equal to 1.

The paper is organized as follows. In section 2 we describe preliminary background information. The main results are presented in section 3 where we derive our bounds and perform an analyzis about the behavior of them as N grows infinitely. In section 4 we describe some numerical results involving Vandermonde matrices related to signal processing applications. Finally, we present some conclusions in section 5.

2. An estimate for $\|\mathbf{W}_\mathbf{N}^\dagger\|$ and other preliminaries. In this section we perform a preliminary analysis of the behavior of $\|W_N^\dagger\|$ as a function of N and present a few basic results. In what follows the singular values of a matrix $A \in \mathbb{C}^{M\times N}$ are denoted by $\sigma_i(A)$ and arranged in increasing order, i.e., $\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_q(A)$, where $q = \min(M, N)$. The spectrum of $A \in \mathbb{C}^{N\times N}$ is denoted by $\lambda(A)$. Also, $\|A\|$ and $\|A\|_F$ denote the spectral and the Frobenius norm of A, respectively; A^* denotes the transpose conjugate of A.

The following result, whose proof follows from the Rayleigh characterization of $\sigma_n^2(W_N)$ as an eigenvalue of $W_N W_N^*$, provides information about the behavior of $\|W_N^\dagger\|$ as a function of N.

Theorem 1. *Let W_N be a Vandermonde matrix as in* (1.1) *but with distinct nodes in the complex plane. Then for $N > n$, $\|W_N^\dagger\|$ decreases monotonically with N and this decrease is strict along the subsequence $\{n, 2n, \ldots\}$. Moreover, if $N = p\cdot n$, $p \geq 1$, then*

$$\|W_{p\cdot n}^\dagger\| \leq \frac{\|W_n^{-1}\|}{\sqrt{1 + \beta^{2n} + \beta^{4n} + \cdots + \beta^{2(p-1)n}}}, \tag{2.1}$$

where $\beta = \min|z_j|$, $j = 1, \ldots, n$.

This preliminary result suggests that, since $\|W_N^\dagger\|$ decreases monotonically with N, it is not unreasonable to expect improvements of $\kappa_2(W_N)$ as N increases, though the behavior of $\kappa_2(W_N)$ as a function of N may be not always monotonic, as illustrated in section 4. We also observe that the bound (2.1) as a function of N may be very pessimistic if W_n is very ill conditioned. In spite of this, it can be very useful for theoretical purposes as we illustrate below.

Theorem 2. *Let $\widehat{f}_N$ be the minimum* 2*-norm solution of the underdetermined system*

$$W_N f = Z^N e,$$

where $Z = \text{diag}(z_1, \ldots, z_n)$ and e is the vector in $\mathbb{R}^n$ of all ones. Suppose that all nodes z_j satisfy either $|z_j| = 1$ or $|z_j| < 1$; then, whenever $N \to \infty$ we have

$$\|\widehat{f}_N\| \to 0.$$

Proof. It is sufficient to note that $\|\widehat{f}_N\| = \|W_N^\dagger Z^N e\| \leq \|W_N^\dagger\|\sqrt{n}\alpha^N$, where $\alpha = \max_j |z_j|$, $j = 1, 2, \ldots, n$, and then to use Theorem 1. ∎

The vector $\widehat{f}_N$ we have just described appears in connection with linear prediction problems and autoregressive modeling of time series. It is currently referred to as *vector of predictor parameters*; see Bazán and Bezerra [2] or Cybenko [6] for details. In our context, it predicts the last column of W_{N+1}. This fact yields the following interesting relationship:

$$ZW_N = W_N \mathcal{F}_N, \tag{2.2}$$

where $\mathcal{F}_N$ is an $N \times N$ companion matrix of the form

$$\mathcal{F}_N = [e_2, e_3, \ldots, e_{N-1}, \widehat{f}_N],$$

in which e_i denotes the ith canonical vector in $\mathbb{R}^N$.

Define F_N by

$$F_N = V_N^* \mathcal{F}_N V_N, \tag{2.3}$$

where

$$V_N = W_N^*(W_N W_N^*)^{-1/2}.$$

Note that F_N is well defined for $W_N W_N^*$ is positive definite. Then the following result is obtained.

THEOREM 3. *The matrix F_N introduced in* (2.3) *has an eigenvalue decomposition given by*

$$F_N = Q_N Z Q_N^{-1}, \tag{2.4}$$

where $Q_N = (W_N W_N^*)^{-1/2}$. Furthermore, this eigenvector matrix satisfies $\kappa_2(Q_N) = \kappa_2(W_N)$.

Proof. The proof is straightforward. ∎

Much of our results concerning bounds for $\kappa_2(W_N)$ depend on the eigenvalue and singular value spectra of F_N. The following theorem characterizes the singular spectrum of F_N.

THEOREM 4. *Let F_N be as in* (2.3)*. Then its singular spectrum is described by*

$$\begin{aligned} \sigma_1^2(F_N) &= \frac{2 + \|\widehat{f}_N\|^2 - \|p_1\|^2 + \sqrt{(\|\widehat{f}_N\|^2 + \|p_1\|^2)^2 - 4|f_1|^2}}{2}, \\ \sigma_j^2(F_N) &= 1, \quad j = 2, n-1, \\ \sigma_n^2(F_N) &= \frac{2 + \|\widehat{f}_N\|^2 - \|p_1\|^2 - \sqrt{(\|\widehat{f}_N\|^2 + \|p_1\|^2)^2 - 4|f_1|^2}}{2}, \end{aligned} \tag{2.5}$$

where p_1 is the first column of $\mathcal{P}_N$, the orthogonal projector onto $\mathcal{R}(W_N^*)$, and f_1 the first component of the vector $\widehat{f}_N$ introduced in Theorem 2.

Proof. Using (2.4) we have that

$$F_N F_N^* = Q_N Z Q_N^{-2} Z^* Q_N = Q_N^* Z W_N W_N^* Z^* Q_N = V_N^* \mathcal{F}_N \mathcal{F}_N^* V_N,$$

where the last equality is because of (2.2). Since $\mathcal{F}_N$ is a companion matrix, this can be now rewritten as

$$F_N F_N^* = I + xx^* - yy^*, \tag{2.6}$$

where $x = V_N^* \widehat{f}_N$ and $y = V_N^* e_1$. Hence, it can be proved (Bazán and Toint [3]) that $\lambda(F_N F_N^*)$ is formed by $n-2$ eigenvalues equal to 1 and the remaining ones are given by

$$\begin{aligned} \gamma_1 &= \frac{2 + \|x\|^2 - \|y\|^2 + \sqrt{(\|x\|^2 + \|y\|^2)^2 - 4(x^*y)^2}}{2}, \\ \gamma_2 &= \frac{2 + \|x\|^2 - \|y\|^2 - \sqrt{(\|x\|^2 + \|y\|^2)^2 - 4(x^*y)^2}}{2}. \end{aligned}$$

On the other hand, note that the Rayleigh quotient associated with (2.6) ensures that

$$\gamma_2 = \min\{\lambda(F_N F_N^*)\} \leq 1 \leq \max\{\lambda(F_N F_N^*)\} = \gamma_1, \tag{2.7}$$

and these inequalities lead to $\gamma_1 = \sigma_1^2(F_N)$ and $\gamma_2 = \sigma_n^2(F_N)$. To conclude the proof, we first note that the columns of V_N form an orthonormal basis of $\mathcal{R}(W_N^*)$ and therefore $\mathcal{P}_N = V_N V_N^*$. Hence we have that

$$\|y\|^2 = y^* y = e_1^* V_N V_N^* e_1 = e_1^* \mathcal{P}_N^* \mathcal{P}_N e_1 = \|\mathcal{P}_N e_1\|^2,$$

since $\mathcal{P}_N$ is a projector, and that

$$\|x\|^2 = x^* x = \widehat{f}_N^* V_N V_N^* \widehat{f}_N = \widehat{f}_N^* \widehat{f}_N = \|\widehat{f}_N\|^2,$$
$$|x^* y| = |\widehat{f}_N^* V_N V_N^* e_1| = |\widehat{f}_N^* e_1| = |f_1|,$$

since $\widehat{f}_N \in \mathcal{R}(W_N^*)$. This concludes the proof. □

3. Bounds for $\kappa_2(W_N)$. We first recall an important result concerning the sensitivity of eigenvalues to matrix perturbations, which tells us that, for given $A \in \mathbb{C}^{n\times n}$, if one knows a measure of closeness between its singular values and the absolute values of its eigenvalues, then one can always estimate the conditioning of the related eigenvalue problem as a function of this closeness (see, for instance, Smith [13]). A conclusion obtained along these lines is that the closeness of *all* singular values to the absolute values of the corresponding eigenvalues implies a well-conditioned eigenvalue problem (Ruhe [12]). Intuitively, this is because in that case, the spectrum of A behaves *almost* as the spectrum of a normal matrix whose eigenvalue problem is perfectly conditioned.

The bounds that we derive rely then on the fact that the conditioning of the eigenvalue problem related to F_N is governed by $\kappa_2(W_N)$ (Theorem 3) and also on the observation that we can measure the above-mentioned closeness exactly (Theorems 3 and 4). To show this we need to recall some measures for the sensitivity of matrix eigenvalue problems. In fact, let u_j and v_j be left and right eigenvectors of A related to the eigenvalue λ_j and define

$$s_j = \frac{u_j^* v_j}{\|u_j\|\|v_j\|}. \tag{3.1}$$

It is well known that when λ_j is simple, the number $|s_j|^{-1}$ is finite and uniquely determined. This number measures the sensitivity of λ_j to perturbations in A and it is known as *the condition number* of λ_j; see, e.g., Wilkinson [20, p. 314]. The theorem below states some relations involving these condition numbers and an overall condition number for the eigenvalue problem measured by the Frobenius norm.

THEOREM 5. *Suppose $A \in \mathbb{C}^{n\times n}$ has an eigenvalue decomposition $A = X\Lambda X^{-1}$ and assume all eigenvalues λ_j of A are simple. Then the following inequalities hold:*

$$1 \le |s_j^{-1}| \le \left[1 + \frac{D^2(A)}{(n-1)\delta_j^2}\right]^{\frac{n-1}{2}}, \quad 1 \le j \le n, \quad \text{where } \delta_j = \min_{\substack{1\le k\le n \\ k\ne j}} |\lambda_j - \lambda_k|, \tag{3.2}$$

and $D^2(A) = \|A\|_F^2 - \sum_{j=1}^n |\lambda_j|^2$. Furthermore, the condition number $\kappa_F(X) = \|X\|_F \|X^{-1}\|_F$ satisfies

$$\sum_{j=1}^n \frac{1}{|s_j|} \le \kappa_F(X) \le n \cdot \max_j \frac{1}{|s_j|} \cdot \frac{\max_j \|e_j^* X^{-1}\|}{\min_j \|e_j^* X^{-1}\|}. \tag{3.3}$$

Proof. Inequalities (3.2) and the left inequality of (3.3) are consequences of Theorems 5 and 3 by Smith [13], respectively. In order to prove the remaining inequality, we first introduce $u_j^* = e_j^* X^{-1}$ and $v_j = X e_j$. Since these vectors are left and right eigenvectors related to λ_j, using the definition of $\kappa_F(X)$ and (3.1), it follows that

$$k_F^2(X) = \sum_{j=1}^{n} \|Xe_j\|^2 \sum_{j=1}^{n} \|e_j^* X^{-1}\|^2 = \sum_{j=1}^{n} \frac{1}{|s_j|^2} \cdot \frac{1}{\|e_j^* X^{-1}\|^2} \sum_{j=1}^{n} \|e_j^* X^{-1}\|^2.$$

The desired inequality follows then by bounding the terms inside the sums. ∎

Note that if A is normal the bounds for $|s_j^{-1}|$ in (3.2) are sharp since $D^2(A) = 0$. Bounds (3.3) will become sharp if besides normality all right eigenvectors have the same 2-norm. Number $D(A)$ is referred to as *departure from normality* of A (see Henrici [10], or Golub and Van Loan [9, p. 314]). Our results concerning bounds for $\kappa_2(W_N)$ are stated in the following main theorem.

THEOREM 6. *Let W_N be the $n \times N$ Vandermonde matrix with nodes z_j in the unit disk. Define $\alpha = \max_j |z_j|$, $\beta = \min_j |z_j|$, and $\delta = \min_{j,k} |z_j - z_k|$, $j \neq k$. Also, let D_N be the departure from normality of matrix F_N defined in* (2.3)*, that is, $D_N^2 = D^2(F_N) = \|F_N\|_F^2 - (|z_1|^2 + \cdots + |z_n|^2)$. Then, for $N > n \geq 2$, the 2-norm condition number of W_N satisfies*

$$\frac{\sigma_1(F_N)}{\alpha} \leq \kappa_2(W_N) \leq \frac{1}{2}\left(\eta + \sqrt{\eta^2 - 4}\right), \tag{3.4}$$

where $\eta = \rho - n + 2$,

$$\rho = n \left[1 + \frac{D_N^2}{(n-1)\delta^2}\right]^{\frac{n-1}{2}} \phi_N(\alpha, \beta), \tag{3.5}$$

and

$$\phi_N(\alpha, \beta) = \sqrt{\frac{1 + \alpha^2 + \alpha^4 + \cdots + \alpha^{2(N-1)}}{1 + \beta^2 + \beta^4 + \cdots + \beta^{2(N-1)}}}. \tag{3.6}$$

Proof. The left inequality follows after taking 2-norm in both sides of (2.4) and using Theorem 3. To prove the right inequality we first bound $\kappa_F(W_N)$. It suffices to note that $\|e_j^* Q_N^{-1}\| = \|e_j^*(W_N W_N^*)^{1/2}\| = \|e_j^* W_N\|$ and then apply (3.3) and (3.2). This yields

$$\kappa_F(W_N) \leq \rho, \tag{3.7}$$

with ρ as in (3.5). Now, we recall a well-known result involving two condition numbers for a nonsingular $X \in \mathbb{C}^{n \times n}$ (see Smith [13, Theorem 1]) which states that

$$n - 2 + \kappa_2(X) + \kappa_2^{-1}(X) \leq \kappa_F(X).$$

Solving this inequality for $\kappa_2(X)$ and adapting the result to our problem, we obtain

$$\kappa_2(W_N) \leq \frac{1}{2}\left[\kappa_F(W_N) - n + 2 + \sqrt{(\kappa_F(W_N) - n + 2)^2 - 4}\right].$$

The proof concludes by using (3.7) in this inequality. ∎

We now analyze the quality of the upper bound (3.4) as a function of N. Note that the bound depends on three factors: the separation of the nodes z_j in the unit disk, the departure from normality D_N, and $\phi_N(\alpha, \beta)$. As the contribution of $\phi_N(\alpha, \beta)$ is only modest because the nodes lie in the unit disk, it is clear that the quality of the bound mostly depends on the ratio $D_N^2/(n-1)\delta^2$. Thus, moderate bounds will be obtained provided D_N^2 is of the same order of magnitude as $(n-1)\delta^2$ and n is not very large. Therefore, what remains to do is to analyze the behavior of D_N as a function of N. As the following lemma shows, this is possible, since by Theorem 4, we know the singular values of F_N (and hence $\|F_N\|_F$) exactly.

LEMMA 7. *Let D_N be the departure from normality of matrix F_N. Then, for each $N \geq n$,*

$$(n-1)+\frac{\prod_{j=1}^{n}|z_j|^2}{1+\|\widehat{f}_N\|^2}-\sum_{j=1}^{n}|z_j|^2 \leq D_N^2 \leq (n-1)+\|\widehat{f}_N\|^2+\prod_{j=1}^{n}|z_j|^2-\sum_{j=1}^{n}|z_i|^2, \tag{3.8}$$

and therefore

$$D_\infty^2 = \lim_{N\to\infty} D_N^2 = (n-1)+\prod_{j=1}^{n}|z_j|^2-\sum_{j=1}^{n}|z_j|^2.$$

Proof. From the fact that the product of the singular values of a square matrix is equal to the product of the absolute values of the corresponding eigenvalues, we have, by Theorem 4,

$$\sigma_1^2(F_N)\sigma_n^2(F_N) = \prod_{j=1}^{n}|z_j|^2.$$

Hence, since $\|F_N\|_F^2 = \sigma_1^2(F_N) + \cdots + \sigma_n^2(F_N)$, using the definition of D_N^2, we obtain

$$D_N^2 = (n-2)+\sigma_1^2(F_N)+\frac{\prod_{j=1}^{n}|z_j|^2}{\sigma_1^2(F_N)}-\sum_{j=1}^{n}|z_j|^2. \tag{3.9}$$

Since

$$1 \leq \sigma_1^2(F_N) \leq 1+\|\widehat{f}_N\|^2, \tag{3.10}$$

by Theorem 4, inequalities (3.8) follow by applying (3.10) into (3.9). The last part follows after taking the limit as N is going to infinity in both sides of (3.8) and using Theorem 2. ∎

Lemma 7 shows that, while the behavior of D_N^2 as a function of N depends on the speed at which $\|\widehat{f}_N\|^2$ converges to zero as N increases, the size of D_N^2 for N large will ultimately depend on the size of the nodes themselves. We therefore conclude that whenever N is large enough and $|z_j| \approx 1$, the number D_N^2 will be quite small (though D_N may be large for small N). If in addition we assume that the number n is not very large, then W_N should be well conditioned unless the nodes are extremely close to each other.

A comment is needed regarding the speed at which $\|\widehat{f}_N\|^2$ converges to zero. This speed is difficult to estimate because it depends to some extent on the behavior of $\|W_N^\dagger\|$ as a function of N, which ultimately depends on the nodes themselves.

Despite this, the author's experience is that in many applications where the nodes are not much smaller than 1, e.g., when analyzing slow-decaying signals for frequency extraction, moderate values of N are sufficient to ensure small values of $\|\widehat{f}_N\|^2$ (see the examples discussed in [3]). As Vandermonde matrices related to signal processing contain a number of nodes that is typically not very large (genuine applications, in NMR and modal analysis, for example, point out n ranging from 2 to 16; see [17], [18], and [1]), we conclude those matrices should be well conditioned even for moderate values of N.

The upper bound in (3.4) also applies to the case of real nodes. However, we are aware that in this case, the bound could behave quite differently than for complex nodes, since for real nodes close to each other the condition $D_N^2 \approx (n-1)\delta^2$ is difficult to satisfy. All these observations are illustrated by numerical examples in section 4.

We now consider other consequences of Lemma 7. Suppose first that we substitute the upper bound (3.8) in (3.7). This yields the bound

$$\kappa_F(W_N) \leq n \left[1 + \frac{(n-1) + \|\widehat{f}_N\|^2 + \prod_{i=1}^n |z_i|^2 - \sum_{i=1}^n |z_i|^2}{(n-1)\delta^2}\right]^{\frac{n-1}{2}} \phi_N(\alpha, \beta), \tag{3.11}$$

where the factor between brackets decreases with N since $\|\widehat{f}_N\|^2$ does also (see [3, Lemma 6]). Consequently, whenever this factor decreases faster than $\phi_N(\alpha,\beta)$ increases, the bound decreases too. This suggests that in these conditions, it is not unreasonable to expect improvements of the conditioning itself as N increases. We believe this is the most reliable explanation for what one often observes in practice, so far without theoretical explanation (see, e.g., [16] or [4]).

Another consequence of Lemma 7 is that the asymptotic value D_∞^2 now allows us to obtain the asymptotic value of the bounds, at least for two *frequent configurations* of the nodes occurring in practical applications: nodes inside the unit disk and nodes on the unit disk. Before this, however, we prove that the following limit

$$\kappa_{2,W_\infty} = \lim_{N\to\infty} \kappa_2(W_N) \tag{3.12}$$

exists and that it is computable for our two configurations of the nodes. This can be useful if one wishes to evaluate the quality of the bounds.

First, we consider the case where the nodes satisfy $|z_j| < 1$, $j = 1\ldots,n$. Note that in this case, if we let $A_N = W_N W_N^*$ and denote the (j,k) entry of this matrix by $(A_N)_{j,k}$, then

$$(A_N)_{j,k} = 1 + z_j\bar{z}_k + (z_j\bar{z}_k)^2 + \cdots + (z_j\bar{z}_k)^{N-1} = \frac{1-(z_j\bar{z}_k)^N}{1-z_j\bar{z}_k}, \quad 1 \leq j,k \leq n, \tag{3.13}$$

where the bar denotes complex conjugation. Let $C = A_N + B_N$, where B_N is an $n \times n$ matrix whose (j,k) entry is

$$(B_N)_{j,k} = \frac{(z_j\bar{z}_k)^N}{1-z_j\bar{z}_k}.$$

It then follows that A_N, B_N, and C are all positive definite, and thus we have that

$$\sigma_j(C) \geq \sigma_j(A_N), \qquad j = 1,\ldots,n. \tag{3.14}$$

We also note that as $B_N = Z^N C Z^{N*}$ we have $\|B_N\| \leq \alpha^{2N}\|C\|$. Using the definition of C, (3.14), and the last inequality, it is not difficult to prove that

$$(3.15)\qquad (1-\alpha^{2N})\kappa_2(C) \leq \kappa_2^2(W_N) \leq \frac{\kappa_2(C)}{1-\alpha^{2N}\kappa_2(C)} \quad \text{if} \quad \alpha^{2N}\kappa_2(C) < 1.$$

These inequalities show that $\kappa_{2,W_\infty} = \sqrt{\kappa_2(C)}$, which can be computed, e.g., via an eigenvalue solver. Although the estimates (3.15) become sharp as N goes to infinity, they are only partially useful in the finite case: (3.15) holds only for

$$(3.16)\qquad N > -\frac{1}{2}\frac{\ln(\kappa_2(C))}{\ln(\alpha)},$$

which may be very large when $\alpha \approx 1$. If the nodes lie on the unit circle, say, $z_j = \exp(\imath t_j)$, with $\imath = \sqrt{-1}$, observe from (3.13) that $(A_N)_{j,j} = N$, $j = 1, \dots, n$. Using this observation and the fact that

$$(3.17)\qquad |(A_N)_{j,k}| = \left|\frac{\sin(N(t_j - t_k)/2)}{\sin((t_j - t_k)/2)}\right| \leq \frac{2}{|z_j - z_k|} \quad \text{for } j \neq k,$$

which follows from (3.13), by applying the Gerschgorin theorem to A_N it is easy to prove that

$$(3.18)\qquad 1 \leq \kappa_2^2(W_N) \leq \frac{1 + \frac{2n-2}{\delta N}}{1 - \frac{2n-2}{\delta N}} \quad \text{if} \quad N > 2(n-1)/\delta.$$

Despite the fact that these estimates could hold for potentially large values of N when δ is very small, they are sufficiently informative to allow us to conclude that $k_{2,W_\infty} = 1$. Note that the same result can be derived from (3.4) because in this case $\phi_N = 1$ and $D_\infty = 0$. Thus we have proved the following lemma.

LEMMA 8. *Suppose the nodes z_j ($j = 1\colon n$) lie either on or inside the unit disk. Then*

$$\kappa_{2,W_\infty} = \begin{cases} 1 & \text{if } |z_j| = 1, \\ \sqrt{\kappa_2(C)} & \text{if } |z_j| < 1, \end{cases}$$

where C is as above.

We are now ready to describe the asymptotic behavior of the bounds in the case where all nodes lie inside the unit disk. This is done in the following corollary.

COROLLARY 9. *Suppose W_N as before and assume all nodes z_j lie inside the unit disk. Then*

$$(3.19)\qquad \frac{1}{\alpha} \leq \kappa_{2,W_\infty} \leq \frac{1}{2}\left(\breve{\eta} + \sqrt{\breve{\eta}^2 - 4}\right),$$

where

$$(3.20)\qquad \breve{\eta} = n\left[1 + \frac{n - 1 + \prod_{j=1}^n |z_j|^2 - \sum_{j=1}^n |z_j|^2}{(n-1)\delta^2}\right]^{\frac{n-1}{2}} \sqrt{\frac{1-\beta^2}{1-\alpha^2}} - n + 2$$

with α, β, and δ as before.

Proof. It suffices to take the limit when N is going to infinity in both sides of (3.4) and take into account inequalities (3.10) and Lemma 7. ∎

If we now suppose that the nodes are complex with the same magnitude, and if, in addition, we assume that the separation of the nodes is not smaller than the distance of the nodes to the unit circle, then the following simplified version of the upper bound of Corollary 9 is obtained.

COROLLARY 10. *Suppose* $|z_1| = |z_2| = \cdots = |z_n| < 1$. *Also, suppose that* $1-\beta^2 \leq \delta^2$. *Then*

$$\kappa_{2,W_\infty} \leq n\, 2^{\frac{n-1}{2}} - n + 2.$$

Proof. First, notice in (3.6) that $\phi_N(\alpha, \beta) = 1$ for all N since $\alpha = \beta$. Also, since $\beta = |z_i| < 1$, observe that $D_\infty^2 = (n-1)(1-\beta^2) + \beta^{2n} - \beta^2 \leq (n-1)(1-\beta^2)$, for $\beta^{2n} < \beta^2$. Using these observations the upper bound of Corollary 9 reduces to

$$\kappa_{2,W_\infty} \leq n\left(1 + \frac{(1-\beta^2)}{\delta^2}\right)^{\frac{n-1}{2}} - n + 2,$$

which ensures the statement of the corollary since by assumption $(1-\beta^2) \leq \delta^2$. □

As our theoretical results suggest that the numbers $\kappa_2(W_N)$ should be rather small under conditions that are reasonable in several signal processing applications, namely, n small and $|z_j|$ not much smaller than 1, it is natural to ask what one may expect about the behavior of $\kappa_2(W_N)$ if the Vandermonde matrix contains a large number of nodes. This question is very difficult to answer for two reasons. First, as we do not know the speed at which $\|\widehat{f}_N\|^2$ converges to zero, extremely large values of N may be required to achieve "reasonably" small values of D_N^2. Second, even if D_N^2 is small, in the absence of a priori information on the configuration of the nodes, the size of the ratio $D_N^2/(n-1)\delta^2$ (which was shown to govern the quality of upper bound in (3.4)), seems impossible to predict. In spite of this, motivated by the results obtained for the case where the nodes lie on the unit circle, we conjecture that if we assume the nodes close to the unit circle and uniformly spaced, then W_N could be reasonably conditioned, though this could hold for excessively large values of N, as suggested in (3.18).

4. Numerical examples. This section numerically illustrates the behavior of the bounds of Theorem 6 and the remarks made throughout the last section. We analyze Vandermonde matrices arising from frequency estimation and system identification problems where the nodes often satisfy $|z_j| \approx 1$, since they are of the form $z_j = e^{(d_j + \imath\omega_j)\Delta t}$, where $\imath = \sqrt{-1}$, $d_j \leq 0$, and $d_j\Delta t \approx 0$. For each example we show the behavior of the upper bound in (3.4) and the conditioning itself as functions of N, illustrating the role of the departure from normality D_N on that behavior. The departure from normality D_N is computed by using (3.9). The asymptotic values of the upper and lower bounds of Corollary 9, which we denote here by L_∞ and U_∞, respectively, are also shown. All numerical experiments were performed in MATLAB.

We also tried to compare estimates (3.15) with those obtained from Theorem 6, but no illustrative comparison was possible. Whereas the bounds provided by this theorem were sufficiently informative for moderate values of N, this no longer occurred with the estimates (3.15) which required large values of N for its realization, because of the restriction (3.16).

4.1. Example 1: Conditioning of a Vandermonde matrix whose nodes are the poles of a stable time invariant linear system. In this example we

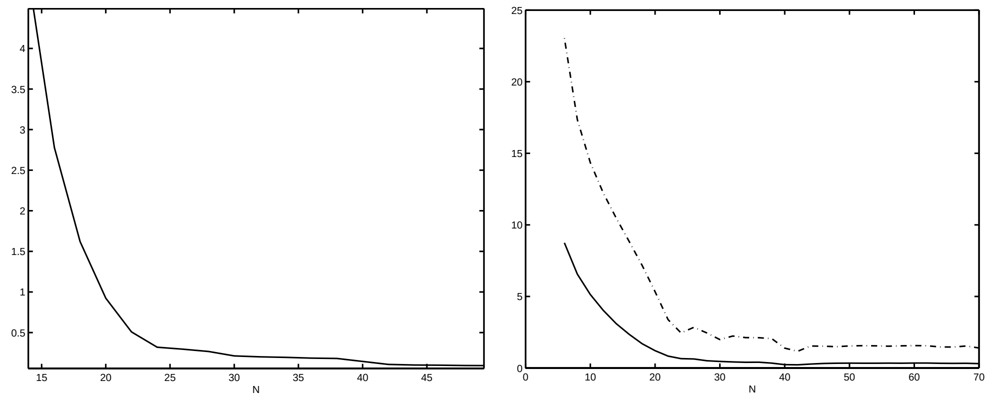


FIG. 1. (a) *Behavior of* $\|\widehat{f}_N\|^2$. (b) *Upper bound of Theorem* 6 *and* $\kappa_2(W_N)$ *on a logarithmic scale* ($\kappa_2(W_N)$: *solid line, upper bound: dashed-dotted line).*

TABLE 1
Nodes of Vandermonde matrix, their absolute values, and corresponding separations.

j	z_j	$\|z_j\|$	δ_j^2
1	$0.9856 \pm 0.1628\imath$	0.9990	0.0794
2	$0.8976 \pm 0.4305\imath$	0.9955	0.0264
3	$0.8127 \pm 0.5690\imath$	0.9921	0.0264

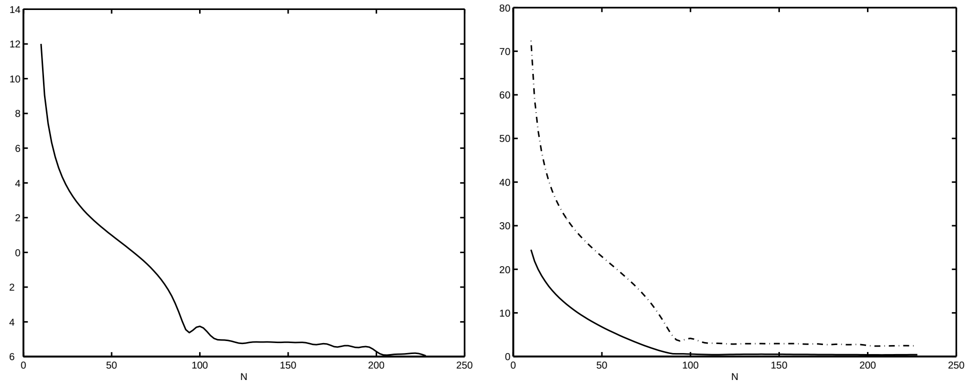
analyze a Vandermonde matrix W_N of order $6 \times N$ (i.e., $n = 6$), whose nodes are the eigenvalues of the matrix

$$A = \begin{pmatrix} 0.9856 & 0.1628 & 0 & 0 & 0 & 0 \\ -0.1628 & 0.9856 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.8976 & 0.4305 & 0 & 0 \\ 0 & 0 & -0.4305 & 0.8976 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.8127 & 0.5690 \\ 0 & 0 & 0 & 0 & -0.5690 & 0.8127 \end{pmatrix}.$$

This matrix concentrates dynamical properties of a simulated discrete time invariant linear system. It is often used in testing system identification algorithms and was taken from [11]. The nodes of W_N, as well as their absolute values and separations δ_j, are illustrated in Table 1.

As in this case all $|z_j|$ are close to 1 and the nodes are not excessively close to each other, the number D_N^2 should be quite small because the fast decrease of $\|\hat{f}_N\|^2$ towards 0, as illustrated in Figure 1(a), even for values of N near to 25. This justifies the relatively small upper bounds for $\kappa_2(W_N)$ for $N \geq 30$ illustrated in Figure 1(b). The logarithmic scale in that figure is because for N near to 6, both $\kappa_2(W_N)$ and the corresponding upper bounds are excessively large. For illustration, while some values of the conditioning itself are $\kappa_2(W_6) = 6.2996 \times 10^3$, $\kappa_2(W_8) = 702.1543$, $\kappa_2(W_{20}) = 3.3559$, for the same values of N the bound decreases from 1.0143×10^{10} down to 206.9487. The situation is even more favorable for $N = 30$, where we obtain $\kappa_2(W_{30}) = 1.5899$ and 7.2880 for the bound.

For this example the limiting values of both the bounds and the conditioning



(a) (b)

FIG. 2. (a) *Behavior of* D_N^2 *on a logarithmic scale.* (b) *Upper bound of Theorem* 6 *and* $\kappa_2(W_N)$ *on a logarithmic scale (*$\kappa_2(W_N)$*: solid line, upper bound: dashed-dotted line.)*

TABLE 2
Nodes of Vandermonde matrix, their absolute values, and corresponding separations.

j	z_j	$\lvert z_j \rvert$	δ_j^2
1	$0.9699 \pm 0.2249\imath$	0.9956	0.0042
2	$0.9532 \pm 0.2931\imath$	0.9972	0.0049
3	$0.9844 \pm 0.1619\imath$	0.9976	0.0031
4	$0.9921 \pm 0.1065\imath$	0.9977	0.0031
5	$0.9972 \pm 0.0485\imath$	0.9989	0.0034

itself are

$$L_\infty = 1.0010, \qquad U_\infty = 12.7454, \qquad \text{and} \quad \kappa_{2,W_\infty} = 2.7927.$$

A comparison of the behavior of bounds of Theorem 6 with that of estimates (3.15) was not possible because the latter hold only for $N \geq 983$.

4.2. Example 2: Conditioning of a Vandermonde matrix arising from a practical application. This example considers a Vandermonde matrix of order $10 \times N$ (i.e., $n = 10$), whose nodes z_j contain information regarding frequencies and decay factors of a vibrating structure. In this case, because the nodes z_j were synthesized from a slow-decaying discrete real signal (see [4] for details), they come in complex conjugate pairs and are close to the unit circle. This is illustrated in Table 2, where the separations of the nodes δ_j are also included.

For this matrix we deduce that the asymptotic values of both the bounds and κ_{2,W_∞} are

$$L_\infty = 1.0016, \qquad U_\infty = 11.8223, \qquad \text{and} \quad \kappa_{2,W_\infty} = 1.7470.$$

Although these asymptotic values are very apparent, it is also important to observe the behavior of the upper bound and the number $\kappa_2(W_N)$ itself as N increases. This is illustrated in Figure 2(b) on a logarithmic scale, from which we see that the upper bound starts to approximate well its asymptotic value $\ln(11.8223) \approx 2.4698$ for values of N near to 90 and that the quality of this approximation is rather good for $N = 200$. We also see that the behavior of the bound for $10 \leq N \leq 90$ is remarkably

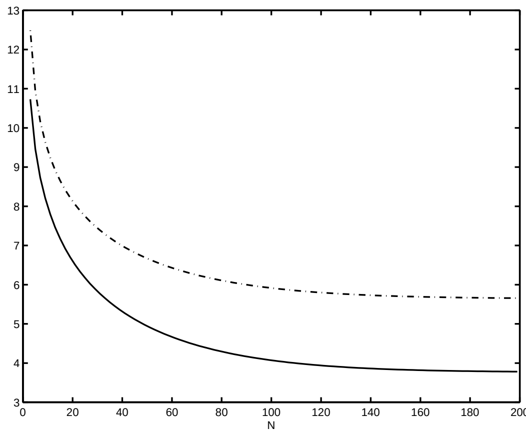


FIG. 3. *Upper bound of Theorem* 6 *(dashed-dotted line) and* $\kappa_2(W_N)$ *(solid line) on a logarithmic scale.*

pessimistic. These two well-distinguished forms of behavior of the bound depend on the behavior of the ratio $D_N^2/(n-1)\delta^2$ in (3.5): while for $10 \leq N \leq 89$, we obtain $0.0327 \leq D_N^2 \leq 1.6348\times10^5$, which indicates that D_N^2 is larger than $(n-1)\delta^2 = 0.0282$, this no longer occurs for $N = 90$, where we obtain $D_N^2 = 0.0189$. The situation is even more favorable when $N = 200$, where $D_N^2 = 0.0036$, a value very close to that of $D_\infty^2 = 0.0012$. The number D_N^2 is displayed as a function of N in Figure 2(a). The behavior of the conditioning itself as a function of N is still more striking: while for $10 \leq N \leq 50$, we obtain $1.3505 \times 10^3 \leq \kappa_2(W_N) \leq 4.3183 \times 10^{10}$, indicating that W_N is ill conditioned for those values of N, that behavior starts to change drastically for $N \geq 50$. A few values illustrate this change: $\kappa_2(W_{50}) = 880.6591$, $\kappa_2(W_{56}) = 264.2212$, and $\kappa_2(W_{60}) = 125.3186$, and for N large, say, $N = 250$, $\kappa_2(W_{250}) = 1.5322$.

Estimates (3.15) hold for $N \geq 383$, and therefore no useful comparison with bounds (3.4) is possible.

4.3. Example 3: Conditioning of a Vandermonde matrix with real nodes. In this example we consider a Vandermonde matrix with real positive numbers as nodes. The nodes arise from an exponential model that simulates fluorescence decay data: $h_k = 2z_1^k + 3z_2^k + 0.1z_3^k$, where $z_1 = e^{-0.05} = 0.9512$, $z_2 = e^{-0.03} = 0.9704$, and $z_3 = e^{-0.02} = 0.9802$ (see [14]). For this case, $(n-1)\delta^2 = 1.9025 \times 10^{-4}$ and $D_\infty^2 = 0.0113$. Hence, as commented in section 4, because the quantity $(n-1)\delta^2$ is always larger than D_N^2, the behavior of the upper bound is rather different from that observed in the case where nodes are complex. The following asympotic values illustrate this observation:

$$L_\infty = 1.02, \qquad U_\infty = 283.2361, \qquad \text{and} \quad \kappa_{2,W_\infty} = 43.2820.$$

The behavior of both the conditioning itself and the corresponding upper bound are illustrated in Figure 3. Estimates (3.15) hold in this case for $N \geq 189$.

5. Conclusions and future work. We have derived lower and upper bounds for the conditioning of rectangular Vandermonde matrices, both depending on the dimension of the problem. The behavior of these bounds for large N allows us to predict that the conditioning itself may be good, under conditions that are reasonable in many applications. Specifically, we have proved that these bounds can be small, provided the dimension of the problem is large enough, with the effect strengthened when the nodes are close to one in absolute value but not excessively close to each other. This was illustrated with numerical examples where we considered Vandermonde matrices with nodes arising from simulated and real applications linked to frequency estimation and parameter identification, in which case the above mentioned conditions are very often satisfied. However, we have not yet investigated the rate at which $\|W_N^\dagger\|$ itself decreases to its limit value.

Perhaps the most important contribution of this work is the qualitative prediction of the improvement of the conditioning of rectangular Vandermonde matrices that is obtained when the dimension is large enough. This information may play a crucial role in the solution of several signal processing applications, in which taking N *large enough* is not a severe restriction. Moreover, given that rank-deficient Hankel matrices are directly related to rectangular Vandermonde matrices, one can determine bounds for the singular values of the former by using the results presented in this paper. This may be of interest in signal subspace approaches for solving frequency estimation problems where the size of the smallest nonzero singular value of Hankel matrices plays a crucial role. This is the object of current research.

Acknowledgments. The author wishes to thank B. Colson, E. Cornelis, and the Department of Mathematics at FUNDP (Namur, Belgium) for providing a cordial environment. Special thanks are due to Ph. L. Toint for helpful discussions and comments regarding the presentation of this paper and to the referees for their suggestions and constructive criticism. Thanks also go to Nick Higham for many suggestions that have improved the presentation of this paper. The author is particularly grateful to one referee for providing estimates (3.15) and (3.18).

REFERENCES

[1] F. S. V. Bazán and C. Bavastri, *An optimized pseudo-inverse algorithm (OPIA) for multi-input multi-output modal parameter identification*, Mechanical Systems and Signal Processings, 10 (1996), pp. 365–380.

[2] F. S. V. Bazán and L. H. Bezerra, *On zero location of predictor polynomials*, Numer. Linear Algebra Appl., 4 (1997), pp. 459–468.

[3] F. S. V. Bazán and Ph. L. Toint, *Singular Values of Predictor Matrices and Signal Eigenvalue Bounds*, Technical report 5, Department of Mathematics, FUNDP, Namur, Belgium, 1998.

[4] F. S. V. Bazán, Ph. L. Toint, and M. C. Zambaldi, *A Conjugate-Gradients Based Method for Harmonic Retrieval Problems that Does Not Use Explicit Signal Subspace Computation*, Technical report 16, Department of Mathematics, FUNDP, Namur, Belgium, 1997.

[5] A. Cordoba, W. Gautschi, and S. Ruscheweyh, *Vandermonde matrices on the circle: Spectral properties and conditioning*, Numer. Math., 57 (1990), pp. 577–591.

[6] G. Cybenko, *Restrictions of normal operators, Padé approximation and autoregressive time series*, SIAM J. Math. Anal., 15 (1984), pp. 753–767.

[7] W. Gautschi, *Norm estimates for inverses of Vandermonde matrices*, Numer. Math., 23 (1975), pp. 337–347.

[8] W. Gautschi and G. Inglese, *Lower bounds for the condition number of Vandermonde matrices*, Numer. Math., 52 (1988), pp. 241–250.

[9] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD, 1996.

[10] P. Henrici, *Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices*, Numer. Math., 4 (1962), pp. 320–329.

[11] J. N. Juang, M. Phan, L. G. Horta, and R. W. Longman, *Identification of observer/Kalman filter Markov parameters: Theory and experiments*, J. Guidance Control Dynam., 16 (1993), pp. 28–36.

[12] A. Ruhe, *On the closeness of eigenvalues and singular values for almost normal matrices*, Linear Algebra Appl., 11 (1975), pp. 87–94.
[13] R. A. Smith, *The condition numbers of the matrix eigenvalue problem*, Numer. Math., 10 (1967), pp. 232–240.
[14] B. Soni and J. Eisenfeld, *System identification of models exhibiting exponential, harmonic and resonant modes*, in Applied Nonlinear Analysis, V. Lakshmikantham, ed., Academic Press, New York, 1979, pp. 555–568.
[15] E. E. Tyrtyshnikov, *How bad are Hankel matrices?*, Numer. Math., 67 (1994), pp. 261–269.
[16] A. Van Der Veen, E. F. Deprettere, and A. Lee Swindlehurst, *Subspace-based signal analysis using singular value decomposition*, Proceedings of the IEEE, 81 (9), 1993, pp. 1277–1309.
[17] S. Van Huffel , *Enhanced resolution based on minimum variance estimation and exponential data modeling*, Signal Processing, 33 (1993), pp. 333–355.
[18] S. Van Huffel, H. Chen, C. Decanniere, and P. Van Hecke, *Algorithm for time-domain NMR data fitting based on total least squares*, J. Magnetic Resonance, A 110 (1994), pp. 228–237.
[19] M. Wei and G. Majda, *A new theoretical approach for Prony's method*, Linear Algebra Appl., 136 (1990), pp. 119–132.
[20] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, UK, 1965.