Singular Values of Predictor Matrices and Signal Eigenvalue Bounds

by F.S.V. Bazán[1] and Ph.L. Toint[2]

Report 98/05 February 5, 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

Current reports available by anonymous ftp from the directory
"pub/reports" on thales.math.fundp.ac.be

Singular Values of Predictor Matrices and Signal Eigenvalue Bounds

F. S. V. Bazán *   Ph. L. Toint

February 5, 1998

Abstract

Predictor matrices often arise in problems of science and engineering where one is interested in predicting future information from previous ones using linear models. The solution of such prediction problems depends on an accurate estimate of a part of the spectrum (*the signal eigenvalues*) of these matrices. In this paper, singular values of predictor matrices are analyzed and formulae for the computation of their singular spectrum are derived. By applying a well-known eigenvalue-singular value inequality to our results, we deduce lower and upper bounds on the modulus of signal eigenvalues. These bounds depend on the dimension of the problem and show that the magnitude of signal eigenvalues is relatively insensitive to small perturbations in the data, provided the signal is slightly damped and the dimension of the problem is large enough. Our theoretical results are illustrated by numerical examples.

1 Introduction

Predictor matrices often arise in connection with problems of science and engineering, where prediction of future from previous information, is of primary interest. In many cases, this prediction is computed using linear models. The nature of the information itself depends on the particular application under study and often has the form of discrete time series, commonly known in engineering as *discrete time signals*. More precisely, given a set of real or complex-valued observations $h_\ell, h_{\ell+1}, \ldots, h_{\ell+N-1}$, a linear prediction model assumes that the future value $h_{\ell+N}$ has the form

$$\bar{f}_1 h_\ell + \bar{f}_1 h_{\ell+1} + \cdots + \bar{f}_N h_{\ell N-1} = h_{\ell+N}, \quad \ell \geq 0. \tag{1.1}$$

In this formulation, N is the *order* of the model and the f's, are complex coefficients known as *predictor parameters*. These coefficients reflect intrinsic properties of the signal such as frequencies, plane waves, dampings, ..., whose estimation from a finite data set $\{h_k\}_{k=0}^{T}$, is an important problem in areas such as system identification and spectral estimation, among others. The linear prediction model is also currently used in other areas such as economy, zoology, ..., see Makhoul

*The research of this author was supported by the CNPq grant 201299/96-8(NV).

(1975) and Box and Jenkins (1970) for details. In this paper , we focus our attention on those applications where the data is real or complex and can be expressed in the form

$$h_k = \sum_{j=1}^{n} r_j e^{(s_j \Delta t)k}, \quad k = 0, 1, \ldots,$$

where r_j and s_j are complex constants such that $s_j \neq s_k$ for $j \neq k$, $s_j = \alpha_j + i\omega_j$, $i = \sqrt{-1}$, $\alpha_j \leq 0$, and Δt is the sampling interval. In this case, the problem is to estimate the parameters r_j, s_j and the number n (*the signal parameters*), from a finite sampling of the observed signal. The classical approach for this problem is relatively simple in the noiseless case: the parameters s_j are extracted from the roots of the so-called *forward predictor polynomial*

$$P_f(t) = \bar{f}_1 + \bar{f}_2 t + \cdots + \bar{f}_N t^{N-1} - t^N, \tag{1.2}$$

whose coefficients $\bar{f}_j$ are estimated by solving the set of linear prediction equations

$$\begin{bmatrix} h_\ell & h_{\ell+1} & \cdots & h_{\ell+N-1} \\ h_{\ell+1} & h_{\ell+2} & \cdots & h_{\ell+N} \\ \vdots & \vdots & \cdots & \vdots \\ h_{\ell+M-1} & h_{\ell+M} & \cdots & h_{\ell+M+N-2} \end{bmatrix} \begin{bmatrix} \bar{f}_1 \\ \bar{f}_2 \\ \vdots \\ \bar{f}_N \end{bmatrix} = \begin{bmatrix} h_{\ell+N} \\ h_{\ell+N+1} \\ \vdots \\ h_{\ell+M+N-1} \end{bmatrix}, \tag{1.3}$$

where $M \geq N \geq n$, and n is the rank of the coefficient matrix. This is justified by the observation that n zeros of $P_f(t)$, known as *signal zeros*, are of the form $z_j = e^{s_j \Delta t}$ (see Wei and Majda, 1990). One could also use the linear prediction equations in the reverse order, replacing (1.1) by

$$\bar{b}_1 h_\ell + \bar{b}_2 h_{\ell+1} + \cdots + \bar{b}_N h_{\ell N-1} = h_{\ell-1}, \quad \ell \geq 1. \tag{1.4}$$

Then the parameters s_j are extracted from the zeros of the *backward predictor polynomials*

$$P_b(t) = \bar{b}_N + \bar{b}_{N-1} t + \cdots + \bar{b}_1 t^{N-1} - t^N, \tag{1.5}$$

whose coefficients are estimated as above. In this case, the signal zeros are $z_j^{-1} = e^{-s_j \Delta t}$ (see, Kumaresan, 1983). However, as only the signal zeros are of interest, one is faced with the problem of separating them from the other $N - n$ zeros, called *extraneous zeros*, which are introduced as a consequence of choosing $N > n$ since n is not known in advance. This separation turns out to be difficult because the extraneous zeros depend on how one chooses the coefficients $\bar{f}_j$ from the infinitely many solutions of the system (1.3). Further details about the zeros of predictor polynomials can be found in Kumaresan (1983), and also in Cybenko (1984), where the problem is examined in the framework of infinite dimensional Hilbert spaces. A more recent approach, where signal zeros are viewed as eigenvalues of predictor matrices (*the signal eigenvalues*), can be found in Bazán and Bezerra (1997). The problem, however, becomes much more difficult when the data is corrupted by noise, since both the rank n and the parameters $\bar{f}_j$ must be estimated from a perturbed linear prediction system, which is generally of full rank. To circumvent this difficulty, the use of high-order predictor models ($N \gg n$) has been proposed and several algorithms have emerged, amongst which the minimum-norm method of Kumaresan and Tufts (1982) and the total least squares-based linear prediction method of Rahman and Yu (1987). Moreover, since methods based on polynomial root-finding are time consuming, new approches based on estimates of the so-called *signal spaces* (the row or column space of the data matrix) have been proposed. In these techniques, one extracts the signal zeros by solving a small $n \times n$ eigenvalue problem (see, for instance, the method proposed in Kung, 1978, known as the HSVD method in Nuclear Magnetic Resonance NMR (see Van Huffel, Chen, Decanniere and Van Hecke , 1994), Schimdt's algorithm, called MUSIC (multiple signal classification, see Schmidt, 1986), the HTLS method by Van Huffel et al., 1994, and OPIA (optimized pseudo-inverse algorithm) by Bazán and Bavastri (1996)). Many references about both polynomial and subspace approaches can be found in Van Der Veen, Deprettere and Lee Swindlehurst (1993).

However, despite the burstling activity in new approaches, little is known about signal eigenvalue sensitivity, an intrinsic component of the problem. It is the purpose of this paper to analyze the magnitude of these eigenvalues, which then allows to show that, under mild conditions, this magnitude is relatively insensitive to small perturbations in the data, at least in many applications. Our result is based on the well-known eigenvalue-singular value relationship

$$\sigma_{min} \leq |\lambda_j| \leq \sigma_{max} \qquad j = 1, 2, \cdots, N, \tag{1.6}$$

(see Golub and Van Loan, 1996, page 318), where σ_{min} and σ_{max} denote the least and largest singular value of any $N \times N$ matrix and λ_j an eigenvalue of the same matrix. We provide analytic formulae for all singular values of a class of predictor matrices and analyze the asymptotic behaviour of related bounds when N tends to infinity.

We first analyze in Section 2 the signal eigenvalues extracted from the spectrum of predictor matrices and their localization. We then derive preliminary results on the singular values of certain predictor matrices. The main results are presented in Section 3 where we give an exact description of the annulus (1.6) for the class of predictor matrices obtained by orthogonal projection: we show that an upper bound on the width of this annulus shrinks as the dimension of the problem increases, and that it asymptotically becomes small provided the signal is slightly damped. Section 4 concludes the paper by presenting some numerical results which illustrate our theoretical analysis.

2 Predictor matrices and basic results

Predictor matrices are defined as follows in the noiseless case. Let $H(\ell)$ be the $M \times N$ Hankel matrix of the system (1.3). We say that the $N \times N$ matrix F is a *forward predictor matrix* if, for all $\ell \geq 0$,

$$H(\ell + 1) = H(\ell)F \tag{2.1}$$

Similarly, B is a *backward predictor matrix* if for $\ell \geq 1$, it satisfies

$$H(\ell - 1) = H(\ell)B. \tag{2.2}$$

We start by presenting some properties of Hankel matrices $H(\ell)$ which appear in these definitions. Notice that $H(\ell)$ can be factored as

$$H(\ell) = VZ^{\ell}RW, \tag{2.3}$$

where $Z = \mathrm{diag}(z_1, \ldots, z_n)$, $z_j = e^{s_j \Delta t}$, $R = \mathrm{diag}(r_1, \ldots, r_n)$, V is an $M \times n$ Vandermonde matrix whose k-th row is the row vector $[z_1^{k-1}\, z_2^{k-1} \cdots z_n^{k-1}]$ and W the transpose of the matrix that consists of the N first rows of V. Hence, we have that (2.3) is a full-rank factorization of $H(\ell)$ and that for all $\ell \geq 0$, $\mathrm{rank}[H(\ell)] = n$. The column space of $H(\ell)$, called *the signal space*, denoted by $\mathcal{R}(H(\ell))$, is spanned by the column vectors of matrix V, while the row space of $H(\ell)$, $\mathcal{R}(H(\ell)^*)$, called the *short signal subspace* (De Moor, 1993), is spaned by the rows of W. Here the symbol * stands for complex conjugate transpose. From this decomposition also follows that $\mathcal{N}(H(\ell))$, the null space of $H(\ell)$, is equal to the null space of W, and that therefore

$$\mathcal{R}[H(\ell)^*] = [\mathcal{N}(W)]^{\perp}. \tag{2.4}$$

Furthermore, if $\mathcal{P}$ denotes the orthogonal projector onto $\mathcal{R}(H(\ell)^*)$, then we have that

$$\mathcal{P} = H(l)^{\dagger}H(l) = W^{\dagger}W = W^*W^{\dagger *}, \tag{2.5}$$

where † stands for the Moore-Penrose pseudo-inverse. The reader is referred to Bjork (1996) for details on projections and pseudo-inverses. The structure of W implies the identity

$$WW^{\dagger} = W^{\dagger *}W^* = I, \tag{2.6}$$

where I denotes the $n \times n$ identity matrix.

We now observe that (2.1) and (2.2) have infinitely many solutions because $H(\ell)$ is rank deficient. The solutions set of (2.1), S_F, is

$$S_F = \{F \mid F = F_M + (I - \mathcal{P})X, \quad X \in R^{N \times N}\} \tag{2.7}$$

where F_M denotes the solution of minimum 2-norm of (2.1). Similarly, we have that the set of backward predictor matrices is

$$S_B = \{B \mid B = B_M + (I - \mathcal{P})X, \quad X \in R^{N \times N}\}, \tag{2.8}$$

where B_M denotes the solution of minimum 2-norm of (2.2). We observe that the signal eigenvalues can be extracted from any forward or backward predictor matrix for, if one substitutes F in (2.1), one has that,

$$WF = ZW, \tag{2.9}$$

which shows that the rows of W are left eigenvectors of F corresponding to the signal eigenvalues. In the same way, (2.2) implies that

$$WB = Z^{-1}W, \tag{2.10}$$

and thus the rows of W are left eigenvectors of B associated with the reciprocal of the signal eigenvalues, $\{z_1^{-1}, z_2^{-1}, \ldots, z_n^{-1}\}$. However, if signal eigenvalues are independent of which predictor matrix is chosen in the class, this is not the case for the extra $N-n$ eigenvalues (*the extraneous eigenvalues*). It turns out that a analysis of the localization of these extraneous eigenvalues is possible, provided we restrict ourselves to a suitable class of predictor matrices. We focus here on two interesting choices: the matrix F_M (or B_M) and minimal norm predictor matrices with companion structure. This last class covers, if prediction is carried out in the forward direction, predictor matrices of the form

$$F = [e_2\, e_3\, \cdots\, e_N\, f],$$

where e_j is the j-th vector of the canonical basis and the N-th column vector, $f = [f_1\, f_2 \cdots f_N]^T$, is the solution of the linear system

$$H(\ell)\, f = H(\ell+1)e_N$$

of minimal ℓ_2-norm Notice that, because of (2.3), this is equivalent to

$$W f = Z^N e, \tag{2.11}$$

where e is the vector of all ones. If prediction is carried out in the reverse order instead, backward companion predictor matrices have the form

$$B = [b\, e_1\, e_2\, \cdots\, e_{N-1}],$$

where the first column vector, $b = [b_1\, b_2 \cdots b_N]^T$, is the solution of the system

$$H(\ell)\, b = H(\ell-1)e_1, \quad \ell \geq 1$$

of minimal ℓ_2-norm. This last system is equivalent to

$$W b = Z^{-1} e. \tag{2.12}$$

We are now ready to state our first result concerning the localization of the eigenvalues of predictor matrices.

Theorem 1 *Let F_M and B_M be the matrices defined in (2.7) and (2.8),respectively. Also let F and B be the forward and backward minimal norm companion predictor matrices. Then, provided $N > n$, we have that*

(a) any predictor matrix with companion structure is nonsingular,

(b) the extraneous eigenvalues of F are the complex conjugates of those of B,

(c)

$$|\hat{\lambda}_j(\cdot)| < 1, \quad j = 1, \ldots N-n$$

where $\hat{\lambda}(\cdot)$ denotes an extraneous eigenvalue and $(\cdot)$ any of the matrices F_M, B_M, F or B.

Proof. Let $V = [V_1\, V_2]$ be an $N \times N$ unitary matrix such that the columns of V_1 form an orthonormal basis of $\mathcal{R}(H(\ell)^*)$ and let F be any forward companion predictor matrix. Consider the unitary transformation

$$V^*FV = \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} F\,[V_1\, V_2] = \begin{bmatrix} V_1^*FV_1 & V_1^*FV_2 \\ V_2^*FV_1 & V_2^*FV_2 \end{bmatrix},$$

and observe that, using (2.5) and (2.9),

$$V_1^*FV_2 = V_1^*V_1V_1^*FV_2 = V_1^*\mathcal{P}FV_2 = V_1^*W^\dagger WFV_2 = V_1^*W^\dagger ZWV_2 = 0,$$

since $\mathcal{R}(V_2) = \mathcal{N}(W)$ because of (2.4). We have then that the spectrum of F, $\lambda(F)$, is

$$\lambda(F) = \lambda(V_1^*FV_1) \cup \lambda(V_2^*FV_2). \tag{2.13}$$

On the other hand, notice that, $\mathcal{P} = V_1V_1^*$ and $\mathcal{Q} = I - \mathcal{P} = V_2V_2^*$, and also observe that, if we write these matrices in partitioned form as $\mathcal{P} = [p_1\; p_2\; \cdots\; p_N]$ and $\mathcal{Q} = [q_1\; q_2 \cdots\; q_N]$, we have that

$$V_1V_1^*F = [p_2 \cdots\; p_{N-1}\; \mathcal{P}f] \;\text{ and }\; V_2V_2^*F = [q_2 \cdots\; q_{N-1}\; \mathcal{Q}f],$$

because F is a companion matrix. This implies that

$$\text{rank}(V_1V_1^*F) = n \;\text{ and }\; \text{rank}(V_2V_2^*F) = N - n, \tag{2.14}$$

since at least n columns of $V_1V_1^*F$ ($N - n$ of $V_2V_2^*F$, respectively) are linearly independent.

However, since discarding zero eigenvalues,

$$\lambda(V_1^*FV_1) = \lambda(V_1V_1^*F) \;\text{ and }\; \lambda(V_2^*FV_2) = \lambda(V_2V_2^*F), \tag{2.15}$$

(see Horn and Johnson (1985), Theorem 1.3.20), we deduce from (2.13) that the spectrum of F is formed by N non-zero eigenvalues and thus F is nonsingular. One can similarly check that any backward companion predictor matrix is nonsingular, and part (a) of the theorem holds. The proof of (b) and that part of c) involving companion matrices can be found in Bazán and Bezerra (1997). We now prove that (c) for F_M and B_M. In order to see that the extraneous eigenvalues of F_M fall inside the unit circle, notice that, since $F_M = \mathcal{P}F$ because of (2.7), we have that

$$\lambda(F_M) = \{z_1, z_2, \cdots, z_n\} \cup \{0\},$$

since $\text{rank}(F_M) = n$ by (2.14). Therefore, part (c) of the theorem holds for F_M. A similar argument can be applied for B_M, which concludes the proof. □

As this theorem describes completely the locations of all eigenvalues of the predictor matrices F_M, B_M, F and B, what remains to do is to determine their singular values. We first start with a technical lemma that enables us to compute the singular spectrum of companion predictor matrices. The determination of the singular spectrum of predictor matrices of minimum 2-norm is slightly more involved and is postponed to the next section.

Lemma 2 *Let a_1, a_2, b_1 and b_2 be vectors in $\mathbb{C}^N$, with $N > 2$, such that at least one of the inner products $b_1^* a_1$ or $b_2^* a_2$ is different of -1. Suppose that the rank-two modification of the identity given by $I + a_1 b_1^* + a_2 b_2^*$ is non singular. Then we have that,*

$$\det(I + a_1 b_1^* + a_2 b_2^*) = 1 + b_1^* a_1 + b_2^* a_2 + b_1^* a_1 b_2^* a_2 - b_2^* a_1 b_1^* a_2, \tag{2.16}$$

and that, the associated characteristic polynomial is

$$p(\lambda) = (1-\lambda)^{N-2}[\lambda^2 - (2 + b_1^* a_1 + b_2^* a_2)\lambda + 1 + b_1^* a_1 + b_2^* a_2 + b_2^* a_2 b_1^* a_1 - b_2^* a_1 b_1^* a_2]. \tag{2.17}$$

Proof. We assume, without loss of generality, that $b_1^* a_1 \neq -1$. Then, from the Sherman-Morrison theorem, we have that $I + a_1 b_1^*$ is non singular since $\det(I + a_1 b_1^*) = 1 + b_1^* a_1$ is non-zero by assumption. Hence, using properties of the determinant, we have that

$$\begin{aligned} \det(I + a_1 b_1^* + a_2 b_2^*) &= \det(I + a_1 b_1^*) \det(I + (I + a_1 b_1^*)^{-1} a_2 b_2^*)) \\ &= (1 + b_1^* a_1)(1 + b_2^* (I + a_1 b_1^*)^{-1} a_2), \end{aligned}$$

and the first part of the lemma follows after applying the Sherman-Morrison formula again to the last right-hand side. On the other hand, given that

$$p(\lambda) = \det(I + a_1 b_1^* + a_2 b_2^* - \lambda I) = \det((1-\lambda) I + a_1 b_1^* + a_2 b_2^*),$$

since $p(1) = 0$ and $N > 2$, we have that $\lambda = 1$ is an eigenvalue of $I + a_1 b_1^* + a_2 b_2^*$. If $\lambda \neq 1$,

$$p(\lambda) = (1-\lambda)^N \det(I + (1-\lambda)^{-1} a_1 b_1^* + (1-\lambda)^{-1} a_2 b_2^*),$$

and the second part of the lemma is then obtained by applying (2.16) in this equation. □

Thus it suffices to extract the eigenvalues associated with a quadratic polynomial in (2.17) to obtain the eigenvalues of the perturbed matrix, since the remaining ones are equal to one. We illustrate this by considering the problem of computing the singular spectrum of the backward companion matrix B introduced above. In fact, since the singular values of B can be computed from the eigenvalues of BB^*, we observe that

$$BB^* = [b \; e_1 \; \dots \; e_{N-1}] \begin{bmatrix} b^* \\ e_1^* \\ \vdots \\ e_{N-1}^* \end{bmatrix} = bb^* + e_1 e_1^* + \cdots + e_{N-1} e_{N-1}^*,$$

where b is the minimum 2-norm solution of (2.12), and hence that

$$BB^* = I + bb^* - e_N e_N^*.$$

By applying Lemma 2 to BB^*, with $a_1 = b_1 = b$ and $a_2 = -b_2 = e_N$, we find that the characteristic polynomial of BB^* is $p(\lambda) = (1-\lambda)^{N-2}[\lambda^2 - \lambda(1 + \|b\|^2) + 4|b^* e_N|^2]$. Hence, we have that the singular spectrum of B, $\sigma(B)$, is of the form

$$\sigma(B) = \{\sigma_1(B), 1, 1, \dots, 1, \sigma_N(B)\},$$

where

$$\sigma_1^2(B), \sigma_N^2(B) = \frac{||b||^2 + 1 \pm \sqrt{(||b||^2+1)^2 - 4|e_N^* b|^2)}}{2}, \tag{2.18}$$

with b_N the N-th component of the predictor coefficients. Since the form of the characteristic polynomial implies that $\sigma_1^2(B) + \sigma_N^2(B) = ||b||^2 + 1$ and since (2.18) ensures that $\sigma_N^2(B) \leq 1$, we also obtain that, as expected, $\sigma_N(B) \leq 1 \leq \sigma_1(B)$. This result is not new, (see for instance Kittaneh, 1995), but the authors are not aware of a proof along the lines developed here. We now use it to obtain an important eigenvalue bound. Since for each signal eigenvalue λ we have that,

$$\sigma_N(B) \leq |\lambda| \leq \sigma_1(B) \leq \sqrt{\frac{||b||^2 + 1 + \sqrt{(||b||^2+1)^2 - 4|e_N^* b|^2)}}{2}} \leq \sqrt{1 + ||b||^2}, \tag{2.19}$$

an interesting upper bound can be immediately derived provided $||b||^2$ is small enough (remember in this case $|\lambda| \geq 1$). In our context, as demonstrated by Bazán, Toint and Zambaldi (1997), it is fortunate that $||b||^2 \approx 0$ in many practical applications, provided the dimension of the problem is sufficiently large. If this is true, then the form of the upper bound indicates that it could be rather tight. Hence, this preliminary analysis suggests that reliable bounds could be obtained provided they only depend on quantities similar to the right-hand side of (2.19). Unfortunately, no lower bound of interest can be obtained from the left inequality because $\sigma_N(B) \approx 0$, which motivates our search for a better lower bound.

3 Predictor matrices obtained by orthogonal projection and signal eigenvalue bounds

In spite of the promising quality of the above upper bound, the link of the signal eigenvalues with the solution of a "large eigenvalue problem" seems to generate a new inconvenience, in that it appears to require that the prediction matrix is sufficiently large. In this section, we shall show that this can be circumvented provided signal eigenvalue bounds are derived by using the singular spectrum of predictor matrices *obtained via orthogonal projections.* We say that the $n \times n$ matrix $F_{\mathcal{P}}$, is a forward predictor matrix obtained by orthogonal projection if, it is of the form

$$F_{\mathcal{P}} = V_1^* F V_1,$$

where V_1 denotes any $N \times n$ matrix with orthonormal columns that span $\mathcal{R}(H(\ell)^*)$. The motivation for this definition relies on the fact that the spectrum of $F_{\mathcal{P}}$ only contains the n signal eigenvalues, since, as we saw in (2.15), $\lambda(F_{\mathcal{P}}) = \lambda(\mathcal{P}\mathcal{F})$ when zero eigenvalues are discarded. Similarly, we say that the $n \times n$ matrix $B_{\mathcal{P}}$, is a backward predictor matrix obtained by projection if, it is of the form

$$B_{\mathcal{P}} = V_1^* B V_1.$$

The spectrum of $B_{\mathcal{P}}$ then consists of the reciprocal of the signal eigenvalues. Thus, if signal eigenvalue bounds are derived from the singular spectrum of these matrices, the signal eigenvalues are related to a small $n \times n$ eigenvalue problem. The purpose of this section is therefore to develop a singular value analysis of these matrices and to analyze the corresponding signal eigenvalue bounds.

Theorem 3 *Suppose B_M is the backward predictor matrix of minimum 2-norm, then*

$$\sigma(B_M) = \sigma(B_{\mathcal{P}}) \cup \{0\},$$

where

$$\begin{aligned} \sigma_1^2(B_{\mathcal{P}}) &= \frac{2 + \|b\|^2 - \|p_N\|^2 + \sqrt{(\|b\|^2 + \|p_N\|^2)^2 - 4|e_N^* b|^2}}{2}, \\ \sigma_i(B_{\mathcal{P}}) &= 1, \quad (i = 2, \dots, n-1), \\ \sigma_n^2(B_{\mathcal{P}}) &= \frac{2 + \|b\|^2 - \|p_N\|^2 - \sqrt{(\|b\|^2 + \|p_N\|^2)^2 - 4|e_N^* b|^2}}{2}, \end{aligned} \tag{3.20}$$

where b and p_N are the first and the last column vectors of B and the projector $\mathcal{P}$, respectively.

Proof. From (2.8) we have that $B_M = \mathcal{P}B = V_1 V_1^* B$. Hence,

$$B_M^* B_M = B^* V_1 V_1^* V_1 V_1^* B = (V_1^* B)^* (V_1^* B),$$

and therefore

$$\sigma(B_M) = \sigma(V_1^* B). \tag{3.21}$$

On the other hand, if we introduce $A = V_1^* B$, then $B_{\mathcal{P}} = A V_1$ and

$$B_{\mathcal{P}} B_{\mathcal{P}}^* = A V_1 V_1^* A^* = A \mathcal{P} A^*. \tag{3.22}$$

But, as $\mathcal{P} = V_1 V_1^* = W^\dagger W = W^* W^{\dagger *}$ by (2.5), then

$$A^* = B^* V_1 = B^* \mathcal{P} V_1 = B^* W^* W^{\dagger *} V_1 = W^* Z^{-*} W^{\dagger *} V_1,$$

where we used (2.10) to obtain the last equality. Hence, since $W^{\dagger *} W^* = I$ by (2.6), we have, using (2.10) again, that

$$\mathcal{P} A^* = W^* W^{\dagger *} W^* Z^{-*} W^{\dagger *} V_1 = B^* W^* W^{\dagger *} V_1 = B^* V_1 = A^*. \tag{3.23}$$

Using this property in (3.22), we deduce that the singular values of $B_{\mathcal{P}}$ are the singular values of A, and thus the first part of the theorem follows from (3.21). In order to prove the second statement, observe that

$$AA^* = [V_1^* b, V_1^* e_1, \dots, V_1^* e_{N-1}] \begin{bmatrix} b^* V_1 \\ e_1^* V_2 \\ \vdots \\ e_{N-1}^* V_1 \end{bmatrix} = V_1^* b b^* V_1 + V_1^* e_1 e_1^* V_1 + \cdots + V_1^* e_{N-1} e_{N-1}^* V_1,$$

can be rewritten as

$$AA^* = I + xx^* - yy^*,$$

where $x = V_1^* b$, and $y = V_1^* e_N$. By applying Lemma 2 to AA^* with $a_1 = b_1 = x$ and $a_2 = -b_2 = y$, we obtain that the spectrum of this matrix is formed by $n-2$ eigenvalues equal to 1, and the remaining ones, obtained from the roots of the polynomial in (2.17), are given by

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

Now, observe that

$$\|p_N\| = \|V_1 V_1^* e_N\| = \|V_1(V_1^* e_N)\| = \|V_1 y\| = \|y\|,$$

since V_1 is an isometry, and that

$$\|x\|^2 = b^* V_1 V_1^* b = b^* \mathcal{P} b = b^* b, = \|b\|^2,$$
$$|x^* y| = |b^* V_1 V_1^* e_N| = |(b^* \mathcal{P}) e_N| = |b^* e_N|$$

since $b \in \mathcal{R}[H(\ell)^*]$. We now observe that the largest value of the above roots, λ_1, say, is larger than one because the eigenvalues of B are larger than one in modulus. Assume now that $\lambda_2 > 1$. Then we obtain from its definition that

$$\left(\|x\|^2 - \|y\|^2\right)^2 > (\|x\|^2 + \|y\|^2)^2 - 4(x^* y)^2,$$

which can be simplified to

$$\|x\|^2 \|y\|^2 > |x^* y|^2.$$

This is impossible as it contradicts the Cauchy-Schwarz inequality, and we therefore deduce that $\lambda_2 \leq 1$. The eigenvalues λ_1 and λ_2 are therefore $\sigma_1^2(B_{\mathcal{P}})$ and $\sigma_n^2(B_{\mathcal{P}})$, respectively, which concludes the proof. □

Thus, we have obtained an exact description of the singular spectrum of predictor matrices obtained by projection and the annulus (1.6) which provides lower and upper bounds for the signal eigenvalues. However, notice that these bounds are not immediately useful because they are derived from the expressions of the singular values given by the last theorem, which depend themselves on the projector $\mathcal{P}$. In order to overcome this difficulty we prove the following result, where we reintroduce the minimum norm solution f of (2.11).

Theorem 4 *Suppose that $\mathcal{A}_N$ is the annulus defined by*

$$\mathcal{A}_N = \left\{ z \in \mathbb{C} \ \middle|\ \frac{1}{\sqrt{1 + \|f\|^2}} \leq |z| \leq \sqrt{1 + \|b\|^2} \right\}, \tag{3.24}$$

where N is the dimension of the backward predictor matrix B, then the eigenvalues of $B_{\mathcal{P}}$ are contained in $\mathcal{A}_N$.

Proof. It suffices to prove that both $\sigma_1(B_{\mathcal{P}})$ and $\sigma_n(B_{\mathcal{P}})$ belong to $\mathcal{A}_N$. In fact, if one uses (3.20), then we have that

$$\sigma_1^2(B_{\mathcal{P}}) = \frac{2 + \|b\|^2 - \|p_N\|^2 + \sqrt{(\|b\|^2 + \|p_N\|^2)^2 - 4|e_N^* b|^2}}{2} \leq 1 + \|b\|^2,$$

which shows that $\sigma_1(B_{\mathcal{P}}) \in \mathcal{A}_N$. In order to prove that $\sigma_n(B_{\mathcal{P}})$ is not smaller than the inner radius of $\mathcal{A}_N$, we first show that $B_{\mathcal{P}} = F_{\mathcal{P}}^{-1}$: using the definitions of both matrices as well as properties (2.5), (2.6), (2.9) and (2.10), we have that

$$B_{\mathcal{P}} F_{\mathcal{P}} = V_1^* B V_1 V_1^* \mathcal{F} V_1 = V_1^* \mathcal{P} B \mathcal{P} \mathcal{F} V_1 = V_1^* W^\dagger W B W^\dagger W \mathcal{F} V_1 = V_1^* W^\dagger Z^{-1} W W^\dagger Z W V_1 = I.$$

One can similarly see that $F_{\mathcal{P}} B_{\mathcal{P}} = I$ and thus the assertion holds. We next observe that this enables us to compute $\sigma_n(B_r)$ as

$$\sigma_n(B_{\mathcal{P}}) = \frac{1}{\sigma_1(F_{\mathcal{P}})}, \tag{3.25}$$

and $\sigma_1(F_{\mathcal{P}})$ can be determined in a way similar to that used for the singular values of the backward predictor matrix. This yields that

$$\sigma_1(F_{\mathcal{P}})^2 = \frac{2 + \|f\|^2 - \|p_1\|^2 + \sqrt{(\|f\|^2 + \|p_1\|^2)^2 - 4|e_1^* f|^2}}{2} \leq 1 + \|f\|^2,$$

where p_1 is the first column of $\mathcal{P}$ and f the minimum norm solution of (2.11). This ensures that the left inequality of (3.24) is satisfied by $\sigma_n(B_{\mathcal{P}})$, which completes the proof.□

We now make the crucial observation that, depending on the dimension of the problem N, the inner and outer radii of $\mathcal{A}_N$ become excellent approximations of the exact values of $\sigma_1(B_{\mathcal{P}})$ and $\sigma_n(B_{\mathcal{P}})$, respectively. This can be seen as follows. Since $e_N^* b$ is the independent coefficient of the characteristic polynomial associated to the companion matrix B (see Bazán and Bezerra (1997) for details), which is not zero by Theorem 1, then

$$|e_N^* b| = \prod_{j=1}^{N-n} |\hat{\lambda}_j| \prod_{j=1}^{n} |\lambda_j|, \tag{3.26}$$

where $\hat{\lambda}_j$ are the so-called extraneous eigenvalues and $\lambda_j = z_j^{-1}$. Hence, as $|\hat{\lambda}_j| < 1$ by Theorem 1, and, since for N large enough, $|e_N^* b|^2 \approx 0$, from (3.20) we have that $\sigma_1^2(B_{\mathcal{P}}) \approx 1 + \|b\|^2$. A similar reasonning on $|e_1^* f|$ gives that $\sigma_1^2(F_{\mathcal{P}}) \approx 1 + \|f\|^2$, and the quality of these approximations improves when N increases.

Before stating our final result, we need two technical lemmas.

Lemma 5 *Define $G_{\mathcal{Q}} = V_2^* B V_2$, where V_2 is an $N \times (N-n)$ matrix whose columns form an orthonormal basis of $\mathcal{N}(H(\ell))$. Then,*

$$\begin{cases} \sigma_i^2(G_{\mathcal{Q}}) & = & 1,\ i = 1, 2 \ldots, N-n-1, \\ \sigma_{N-n}^2(G_{\mathcal{Q}}) & = & 1 - (1 + \|b\|^2)\|q_1\|^2, \end{cases} \tag{3.27}$$

where q_1 is the first column of the orthogonal projector $\mathcal{Q}$ onto $\mathcal{N}(H(\ell))$.

Proof. The proof is in appendix A1.

Notice that, since the spectrum of $G_{\mathcal{Q}}$ is formed by the extraneous eigenvalues of B because of Theorem 1, and, since the product of singular values of any square matrix is equal to the product of the absolute values of their corresponding eigenvalues, this theorem and (3.26) ensure that $|e_N^* b|^2$ finally depends on the behaviour of the smallest singular value of $G_{\mathcal{Q}}$.

Lemma 6 *Let b and f be the first and last column vector of B and F respectively. Then*

$$\frac{1+\|b\|^2}{1+\|f\|^2} = \prod_{j=1}^{n} |z_j|^{-2}$$

for $N > n$. Moreover, both $\|b\|$ and $\|f\|$ decrease monotonocally when N increases.

Proof. The proof is in appendix A2.

We now return to our main objective and continue analyzing the behaviour of the width of $\mathcal{A}_N$ as function of $\|b\|$ and $\|f\|$, and, consequently, of N, we notice that, because of Theorem 6, this reduces to analyzing the norm of the forward coefficients $\|f\|$. But, since

$$\|f\| = \|W^{\dagger}Z^N e\| \leq \|W^{\dagger}\|\sqrt{n}\beta^N, \tag{3.28}$$

by (2.9), where we set $\beta = \max\{|z_i|, \quad i = 1, 2 \ldots, n\}$, it suffices to analyze $\|W^{\dagger}\|$ as function of N. We now choose, for notational convenience, N as a multiple of the number de frequencies, n, i.e., $N = p \times n$, $p > 1$. We also write

$$W = [W_0 \;\; DW_0 \;\; \ldots \;\; D^{p-1}W_0],$$

where W_0 is the $n \times n$ Vandermonde matrix whose i, j entry is z_i^{j-1}, and $D = \mathrm{diag}(z_1^n, z_2^n, \ldots, z_n^n)$. Using these definitions, one can then prove that the smallest singular value of W, $\sigma_n(W)$, satisfies

$$\sigma_n(W) \geq \sigma_n(W_o)\sqrt{\sum_{j=1}^{p} \gamma^{2n(j-1)}}, \tag{3.29}$$

where $\gamma = \min\{|z_i|, \quad i = 1, \ldots, n\}$ (see Bazán et al. (1997), Theorem 1, for details). Hence, we have that

1) if $\alpha_i < 0$, i.e. the signal is damped, from (3.28) we have then that

$$\lim_{N\to\infty} \|f\| = 0, \tag{3.30}$$

because $\beta^N \to 0$ and $\|W^{\dagger}\| = 1/\sigma_n(W)$ is bounded using (3.29) as $\gamma < 1$ ensures that the sum is finite;

2) if the signal is undamped instead, i.e. $\alpha_i = 0$, then (3.29) implies that $\|W^{\dagger}\|$ converges to zero as N tends to infinity because $\beta = \gamma = 1$, and thus once more we obtain the limit (3.30) from (3.28).

But the limit (3.30) and Lemma6 together give that

$$\lim_{N\to\infty}(1 + \|b\|^2) = \prod_{k=1}^{n} |z_k|^{-2}(1 + \lim_{N\to\infty} \|f\|^2) = \prod_{k=1}^{n} |z_k|^{-2}.$$

Summarizing this discussion, we conclude that the second part of Lemma 6, this last limit and (3.24) ensures the following result.

Theorem 7 *The annulus that contains the signal eigenvalues associated to $B_{\mathcal{P}}$ has a monotonically decreasing width, i.e.*

$$\mathcal{A}_{N+1} \subseteq \mathcal{A}_N.$$

Moreover, it is asymptotically described by

$$\mathcal{A}_\infty = \{z \in \mathbb{C} \;/ 1 \le |z| \le \prod_{k=1}^{n} |z_k|^{-1}\}. \tag{3.31}$$

Thus, we have shown that the quality of the signal eigenvalues bounds depends on the speed at which $\|f\|^2$ approaches to zero as the dimension N increases. However, as illustrated in (3.28), this speed depends on the behaviour of $\|W^\dagger\|$ as a function of N, which ultimately depends on the nature of the signal itself. The authors' experience is that in most of practical applications, moderate values of N are sufficient to ensure values of the norms of the predictor coefficients smaller than one (see, for instance, the examples discussed in Bazán et al., 1997).

Now consider the case of slightly damped signals. For such signals, we know that the signal eigenvalues are relatively close to one, which we have shown to imply, for large N, that the width of the annulus (1.6) is small. Since these radii provide excellent approximations for $\sigma_1(B_{\mathcal{P}})$ and $\sigma_n(B_{\mathcal{P}})$, these singular values must be close to each other. Furthemore, the stability of the singular values (see Golub and Van Loan, 1996, Section 8.3.1)guarantees that a small perturbation of the data will not alter this property. This, in turn, implies that the width of the annulus (1.6) remains small, even after a small perturbation. As a consequence, the magnitude of the signal eigenvalues cannot vary by a large amount. We delay the stability analysis of the eigenvalues themselves to a future contribution.

4 Numerical Examples

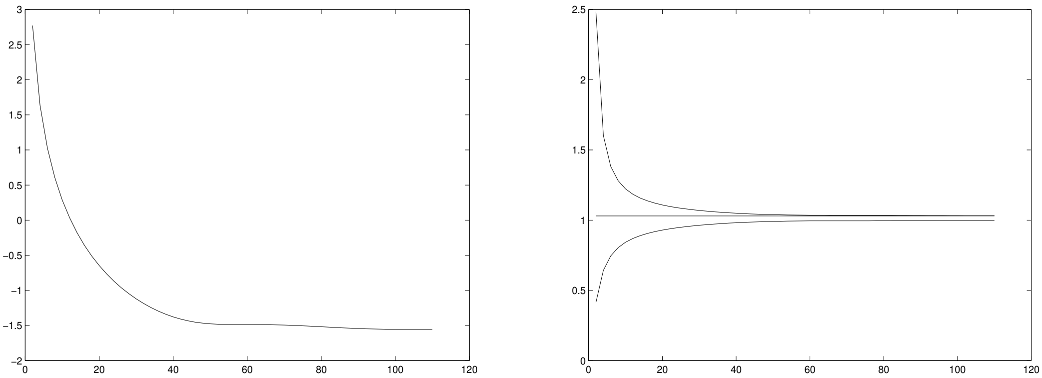
In this section we present the resuls of some computer simulations which illustrate the behaviour of the bounds (3.24) and the role of $\|W^\dagger\|$. We consider two numerical examples. The first is extracted from the specialized literature of the signal analysis field, and the second is a signal synthesized from experimental measurements. For each example, we compute $\|W^\dagger\|$ and the bounds (3.24), $(1+\|f\|^2)^{-\frac{1}{2}}$ and $(1+\|b\|^2)^{\frac{1}{2}}$, for several values of N.

4.1 Bounds for the Signal Eigenvalues of a Synthetic Signal

This example illustrates the bounds associated with the sampled signal defined by

$$h_k = e^{(-0.01+2\pi 0.20i)k} + e^{(-0.02+2\pi 0.22i)k}, \quad k = 0, 1, \ldots,$$

whose signal eigenvalues are $z_1 = e^{(-0.01+2\pi 0.20i)}$ and $z_2 = e^{(-0.02+2\pi 0.22i)}$. In this case, we have that $|z_1| = 0.9900$, $|z_2| = 0.9802$ and consequently the upper bound limit is $\prod_{k=1}^{n} |z_k|^{-1} = 1.035$, (notice that $n = 2$). This signal is often used for testing the capability of algorithms to extract frequencies and decay factors from noisy signals, because the two closely spaced frequencies are easily seen as a single one when additive noise is present (see, for instance, Hua and Sarkar, 1990). The behaviour of the bounds as functions of N is displayed in Figure 1-(b). In Figure 1-(a) we show the behaviour of $\|W^\dagger\|$ on a logarithmic scale. From this figures we see the marked



(a) (b)

Figure 1: (a) $\log(\|W^\dagger\|)$ as function of N. (b) Bounds (3.24) as function of N

monotonic decrease of both the width of the annulus $\mathcal{A}_N$ and $\|W^\dagger\|$ for increasing values of the dimension. In this example, we also verify that for $N \geq 60$, our bounds agree with their limiting values 1 and 1.035 up to two decimal places.

4.2 Bounds for Signal Eigenvalues Synthesized from a Practical Application

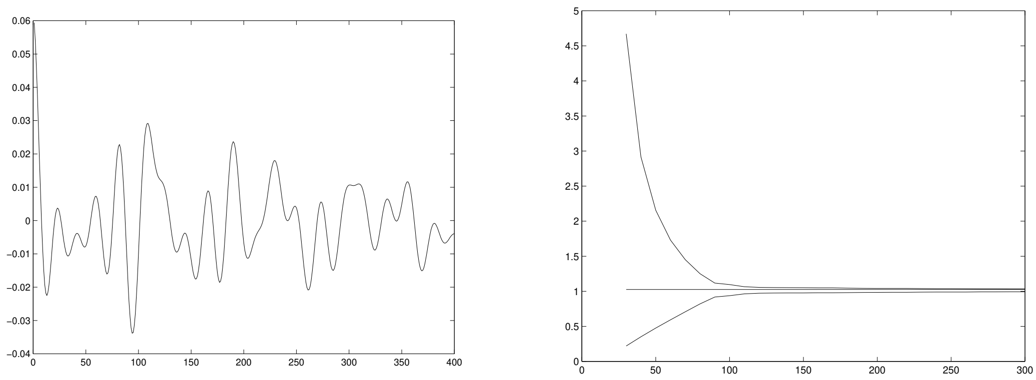
In this example we illustrate the behaviour of the bounds (3.24) for a synthezised signal obtained from experimental measurements of the free response of a real vibratory system. Full information about the procedure used to synthesize the signal can be found in Bazán et al., 1997. For this case, because the signal is real, the signal eigenvalues come in complex conjugate pairs. These eigenvalues are shown in Table 1.

j	z_j	$\|z_j\|$	$\|z_j\|^{-1}$
1	0.9699+ 0.2248i	0.9956	1..0044
2	0.9532+ 0.2931i	0.9972	1.0028
3	0.9844+ 0.1619i	0.9976	1.0024
4	0.9921+ 0.1055i	0.9977	1.0023
5	0.9972+ 0.0585i	0.9989	1.0011

Table 1: Signal Eigenvalues of synthesized signal and corresponding moduli.

The behaviour of the upper and lower bounds (3.24) as funtions of N, which we denote here by L_N and U_N respectively, is displayed for $N \geq 30$ in Figure 2-(b). The rapid decrease of the width of the annulus $\mathcal{A}_N$ is again very apparent.

Notice that, because the absolute values, $|z_j|$, are close to one, the signal is slightly damped (this can also be appreciated by observing the shape of the backward coefficients in Figure



(a) (b)

Figure 2: (a) Backward coefficients $N = 400$. (b) Bounds (3.24) as function of N

2-(a)). Hence, as indicated in the previous section, the decrease of the upper bound to its limiting value could be rather slow. In order to better illustrate the behaviour of the bounds as functions of N, we have computed their distances to their corresponding limits, $L_\infty = 1$, and $U_\infty = \prod_k^{10} |z_k|^{-1} = 1.0262$, respectively. These distances as well as the norms $||W^\dagger||$ are shown in Table 2 for certain values of N. This table also illustrates the decrease of $||W^\dagger||$ as the effect of increasing N. We also note that the bounds agree well with their limits for $N \geq 280$.

N	10	20	30	40	50	60
U_N	415.0088	11.7560	4.6708	2.9155	2.1564	1.7303
$U_N - U_\infty$	413.9826	10.7298	3.6445	1.8893	1.1302	0.7041
L_N	0.0025	0.0873	0.2197	0.3520	0.4759	0.5931
$L_\infty - L_N$	0.9975	0.9127	0.7803	0.6480	0.5241	0.4069
$\|\|W^\dagger\|\|$	$5.4548e+09$	$0.0012e+09$	$1.8228e+04$	$0.0974e+04$	$0.0096e+04$	$0.0014e+04$
N	200	220	240	260	280	300
U_N	1.0411	1.0410	1.0379	1.0367	1.0352	1.0338
$U_N - U_\infty$	0.0149	0.0148	0.0117	0.0105	0.0090	0.0076
L_N	0.9857	0.9858	0.9888	0.9899	0.9913	0.9926
$L_\infty - L_N$	0.0143	0.0142	0.0112	0.0101	0.0087	0.0074
$\|\|W^\dagger\|\|$	0.1127	0.1104	0.1095	0.1055	0.1044	0.1033

Table 2: Behaviour of bounds (3.24) and $||W^\dagger||$ as functions of N.

5 Conclusions and perspectives

We have developed a singular value analysis of certain predictor matrices that enabled us to derive a closed form for their singular spectrum. Using these results, we have derived new lower and upper bounds for the so-called *signal eigenvalues*, both depending of the dimension of the problem. By analyzing the influence of the dimension on these bounds, we have shown that they can become very tight provided the dimension of the problem is sufficiently large. We have also indicated that these bounds are especially useful if one considers slightly damped signals. This was illustrated with numerical examples including the analysis of the new bounds for a signal arising from a practical application.

We may anticipate interesting applications of our results to subspace-approaches for solving the signal parameter retrieval problem, such as HTLS, HSVD, and OPIA. In particular, they seem to open the way for a unified signal eigenvalue perturbation analysis covering these methods, provided they can be shown to depend on specific predictor matrices obtained by projection. This challenging development is the object of ongoing research.

Acknowledgement

The authors wish to thank B. Colson, E. Cornelis and the Departement of Mathematics at FUNDP for providing a cordial environment.

Appendix

A Proof of Lemma 5

Notice that b is the minimum norm solution of (2.12) and therefore $b \in \mathcal{R}(H(\ell)^*)$. From this, we have that

$$G_{\mathcal{Q}} = V_2^* \, B \, V_2 = [0, \, V_2^* e_1, \ldots, V_2^* e_{N-1}] \, V_2,$$

can be rewritten as

$$G_{\mathcal{Q}} = V_2^* \, \mathcal{J} \, V_2 - uv^*, \tag{A.1}$$

where $\mathcal{J}$ is the permutation matrix $\mathcal{J} = [e_N, e_1, \ldots, e_{N-1}]$, and u^* and v^* are respectively, the last and first row of V_2. The proof of the theorem is then based on the computation of the eigenvalues of $G_{\mathcal{Q}}^* G_{\mathcal{Q}}$. We start then by observing that

$$G_{\mathcal{Q}}^* G_{\mathcal{Q}} = G^*G - G^*uv^* - vu^*G + (u^*u)vv^*, \tag{A.2}$$

where we set $G = V_2^* \mathcal{J} V_2$. Analyzing the first term of the right-hand side we see that

$$G^*G = V_2^* \mathcal{J}^* V_2 V_2^* \mathcal{J} V_2 = V_2^* \mathcal{J}^* (I - V_1 V_1^*) \mathcal{J} V_2 = I - V_2^* \mathcal{J}^* V_1 V_1^* \mathcal{J} V_2, \tag{A.3}$$

where the last equality follows from the the orthogonality of $\mathcal{J}$. On the other hand, observe that $\mathcal{J}^* V_1$ can be rewritten as,

$$\mathcal{J}^* V_1 = \begin{bmatrix} w^* \\ V_1^{\uparrow} \end{bmatrix} = \begin{bmatrix} b^* V_1 \\ V_1^{\uparrow} \end{bmatrix} + \begin{bmatrix} w^* - b^* V_1 \\ 0 \end{bmatrix} = A^* + \begin{bmatrix} w^* - b^* V_1 \\ 0 \end{bmatrix}$$

where $V_1^{\uparrow}$ is the matrix of V_1 consisting of all rows excluding the last, w^* is the last row of V_1, and $A^* = BV_1$. Hence, we have that

$$V_2^* \mathcal{J}^* V_1 = V_2^* \begin{bmatrix} w^* - b^* V_1 \\ 0 \end{bmatrix} = V_2^* e_1 (w^* - b^* V_1) = v(w^* - b^* V_1), \tag{A.4}$$

since $V_2^* A^* = 0$ by (3.23). Substituting this relation in (A.3) yields that

$$G^* G = I - (\|w\|^2 - b^* e_N - e_N^* b + \|b\|^2) vv^*. \tag{A.5}$$

We now analyze the second term of the right-hand side of (A.2). Notice that by using $u = V_2^* e_N$ and $v^* = e_1^* V_2$, we have,

$$G^* u = V_2^* \mathcal{J}^* V_2 V_2^* e_N = V_2^* \mathcal{J}^* (I - V_1 V_1^*) e_N = V_2^* \mathcal{J}^* e_N - V_2^* \mathcal{J}^* V_1 V_1^* e_N.$$

But, if one observes that $\mathcal{J}^* e_N = e_1$ and $V_1^* e_N = w$, we have, using (A.4), that

$$G^* u = v - v(w^* - b^* V_1) w = v - v(w^* w - b^* V_1 V_1^* e_N) = v - v(w^* w - b^* e_N),$$

since $b \in \mathcal{R}(H(\ell)^*)$, and hence,

$$G^* u v^* = (1 - \|w\|^2 + b^* e_N) vv^*. \tag{A.6}$$

As a consequence of this derivation, we have that the third term of the right-hand side of (A.2) is

$$v u^* G = (1 - \|w\|^2 + e_N^* b) vv^*. \tag{A.7}$$

Replacing now (A.7), (A.6) and (A.5) into A.2 and taking into account the fact that

$$\|w\|^2 + \|u\|^2 = 1,$$

because $[w^* u^*]$ is the last row of the orthogonal matrix $[V_1 V2]$, we deduce that

$$G_Q^* G_Q = I - (1 + \|b\|^2) vv^*.$$

From this relation, we see that $N - n - 1$ eigenvalues of $G_Q^* G_Q$ are equal to the unity, while the remaining one is $1 - (1 + \|b\|^2)\|v\|^2$. It remains therefore to prove that the norm of v is equal to the norm of the first column vector of the projector $\mathcal{Q}$, which immediately follows from the identity $q_1 = V_2 V_2^* e_1 = V_2 v$.$\square$

B Proof of Lemma 6

We first derive auxiliary results involving the terms of the ratio

$$\frac{1 + \|b\|^2}{1 + \|f\|^2}$$

as functions of N. For this, we consider two consecutive values of N, and use the subscript $_{[N]}$ to denote the dependence of the considered quantities on N. We start by observing that

$$W_{[N+1]}^{\dagger *} W_{[N+1]}^{\dagger} = (W_{[N+1]} W_{[N+1]}^{*})^{-1} \stackrel{\text{def}}{=} A_{[N+1]} \tag{B.8}$$

and that the $W_{[N+1]} W_{[N+1]}^{*} = A_{[N+1]}^{-1}$ is a rank-one modification of $A_{[N]}^{-1} = W_{[N]} W_{[N]}^{*}$, i.e.

$$A_{[N+1]}^{-1} = A_{[N]}^{-1} + Z^N e e^* Z^{N*} = Z A_{[N]}^{-1} Z^* + e e^*,$$

where we used the two representations $W_{[N+1]} = [W_{[N]} \quad Z^N e] = [e \quad Z W_{[N]}]$. Applying the Sherman-Morrison formula to each of these forms, we derive that

$$A_{[N+1]} = A_{[N]} - \frac{A_{[N]} Z^N e e^* Z^{N*} A_{[N]}}{1 + e^* Z^{N*} A_{[N]} Z^N e} = Z^{-*} A_{[N]} Z^{-1} - \frac{Z^{-*} A_{[N]} Z^{-1} e e^* Z^{-*} A_{[N]} Z^{-1}}{1 + e^* Z^{-*} A_{[N]} Z^{-1} e}. \tag{B.9}$$

Now, the projector associated with the value N is given by

$$\mathcal{P}_{[N]} = W_{[N]}^{\dagger} W_{[N]} = W_{[N]}^{\dagger} [e \ \ Z W_{[N-1]}] = [W_{[N]}^{\dagger} e \ \ W_{[N]}^{\dagger} Z W_{[N-1]}]$$

because of (2.5) and the definition of $W_{[N]}$. This yields that

$$\|p_{1,[N]}\| = \|W_{[N]}^{\dagger} e\|.$$

Combining this with (B.8), we obtain that

$$\|p_{1,[N+1]}\|^2 = e^* A_{[N+1]} e.$$

Using this relation in the first equality of (B.9), we obtain

$$\begin{aligned} \|p_{1,[N+1]}\|^2 &= e^* A_{[N]} e - \frac{e^* A_{[N]} Z^N e e^* Z^{N*} A_{[N]} e}{1 + e^* Z^{N*} A_{[N]} Z^N e} \\ &= \|p_{1,[N]}\|^2 - \frac{(e_1^* W_{[N]}^{\dagger} Z^N e)(e^* Z^{N*} W_{[N]}^{\dagger *} e_1)}{1 + e^* Z^{N*} A_{[N]} Z^N e} \\ &= \|p_{1,[N]}\|^2 - \frac{|e_1^* f_{[N]}|^2}{1 + \|f_{[N]}\|^2}, \end{aligned} \tag{B.10}$$

where $f_{[N]} = W_{[N]}^{\dagger} Z^N e$ is the minimum norm solution of the system (2.11). On the other hand, using the equality between the left-hand side and the last right-hand side of (B.9), we have that

$$e^* A_{[N+1]} e = e^* Z^{-*} A_{[N]} Z^{-1} e - \frac{e^* Z^{-*} A_{[N]} Z^{-1} e e^* Z^{-*} A_{[N]} Z^{-1} e}{1 + e^* Z^{-*} A_{[N]} Z^{-1} e}.$$

This is nothing but

$$\|p_{1,[N+1]}\|^2 = \|b_{[N]}\|^2 - \frac{\|b_{[N]}\|^4}{1 + \|b_{[N]}\|^2} = \frac{\|b_{[N]}\|^2}{1 + \|b_{[N]}\|^2},$$

from wich we obtain

$$1 + \|b_{[N]}\|^2 = \frac{1}{1 - \|p_{1,[N+1]}\|^2}. \tag{B.11}$$

We now observe that, using the fact that $p_{1,[N]} = e_1 - q_{1,[N]}$, (B.10) can be rewritten as

$$1 - \|p_{1,[N+1]}\|^2 = \|q_{1,[N]}\|^2 + \frac{|e_1^* f_{[N]}|^2}{1 + \|f_{[N]}\|^2},$$

which, combined with (B.11), gives

$$1 = (1 + \|b_{[N]}\|^2)\|q_{1,[N]}\|^2 + \frac{|e_1^* f_{[N]}|^2}{1 + \|f_{[N]}\|^2}(1 + \|b_{[N]}\|^2),$$

or equivalently, by Theorem 5,

$$\sigma^2_{N-n,[N]}(G_{\mathcal{Q}}) = |e_1^* f_{[N]}|^2 \frac{1 + \|b_{[N]}\|^2}{1 + \|f_{[N]}\|^2}. \tag{B.12}$$

The final part of our proof depends on two important observations. The first is that, as $e_1^* f_{[N]}$, is the independent term of the characteristic polynomial of F, that is non zero because of Theorem 2, it is equal to the product of the eigenvalues of F. That is,

$$|e_1^* f_{[N]}| = \prod_{j=1}^{N-n} |\hat{\lambda}_j| \prod_{j=1}^{n} |z_j|, \tag{B.13}$$

where the $\hat{\lambda}$'s are the extraneous eigenvalues of F, and the z's are the signal eigenvalues. The second is that the extraneous eigenvalues of B are the conjugate of those of F, as proved in Bazán and Bezerra (1997), Theorem 3.2. Hence the product of their modulus is equal to the product of the singular values of $V_2^* B V_2 = G_{\mathcal{Q}}$. Using now Theorem 5, we deduce that

$$\sigma^2_{N-r,[N]}(G_{\mathcal{Q}}) = \prod_{j=1}^{N-n} |\hat{\lambda}_j|^2. \tag{B.14}$$

The first part of the theorem then follows from (B.14), (B.13) and (B.12).

Now, observe that, using the second equality of (B.9) and the definitions of f and b,

$$\begin{aligned} \|f_{[N+1]}\|^2 &= e^* Z^{N+1*} A_{[N+1]} Z^{N+1} e \\ &= e^* Z^{N*} A_{[N]} Z^N e - \frac{e^* Z^{N*} A_{[N]} Z^{-1} e e^* Z^{-*} A_{[N]} Z^N e}{1 + e^* Z^{-*} A_{[N]} Z^{-1} e} \\ &= \|f_{[N]}\|^2 - \frac{|f_{[N]}^* b_{[N]}|^2}{1 + \|b_{[N]}\|^2}, \end{aligned}$$

which shows that $\|f\|$ decreases monotonically with N. The same conclusion then follows for $\|b\|$ because of the first part of the theorem. $\square$

References

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**, 365–380, 1996.

F. S. V. Bazán and L. H. Bezerra. On zero location of predictor polynomials. *Numer. Linear Algebra With Applications*, **4**(6), 459–468, 1997.

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, November 1997.

A. Bjork. *Numerical Methods for least squares problems.* SIAM, Philadelphia, USA, 1996.

G. E. Box and G. M. Jenkins. *Time Series Forecasting and Control.* Holden-Day, San Francisco, 1970.

G. Cybenko. Restrictions of normal operators, padé approximation and autoregressive time series. *SIAM Math. Anal.*, **4**, 753–767, July 1984.

B. De Moor. The singular value decomposition and long and short spaces of noisy matrices. *IEEE Trans. On Signal Proc.*, **41**(9), 2826–2838, September 1993.

G. H. Golub and C. F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore, third edn, 1996.

R. A. Horn and C. R. Johnson. *Matrix analysis.* Cambridge University Press, Cambridge, UK, 1985.

Y. Hua and T. P. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Trans. On Acoust. Speech and Signal Processings*, **ASSP-38**(5), 814–824, May 1990.

Fuadd Kittaneh. Singular values of companion matrices and bouds on zeros of polynomials. *SIAM Journal on Matrix Analysis and Applications*, **16**(1), 333–340, January 1995.

R. Kumaresan. On the zeros of the linear prediction-error filters for deterministic signals. *IEEE Trans. On Acoust. Speech and Signal Processing*, **ASSP-31**(1), 217–221, February 1983.

R. Kumaresan and D. W. Tufts. Estimating the parameters of exponentially damped sinudoids and pole zero-modeling in noise. *IEEE Trans. On Acoust. Speech and Signal Processing*, **ASSP-30**, 833–840, December 1982.

S. Kung. A new identification and model reduction algorithm via singular value decompositions. *in* 'Proceedings of the 12^{th} Asilomar Conference on Circuits Systems and Computers', pp. 705–714, Pacific Grove California, 1978.

J. Makhoul. Linear prediction: A tutorial review. *Proceedings of the IEEE*, **63**(4), 561–580, 1975.

M. A. Rahman and K. B. Yu. Total least squares approach for frequency estimation using linear prediciton. *IEEE Trans. Acoust. Speech Signal process.*, **ASSP-35**, 1440–1454, 1987.

R. O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag*, **AP-34**, 276–281, 1986.

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), 1277–1309, September 1993.

S. Van Huffel, H. Chen, C. Decanniere, and P. Van Hecke . Improved total least squares based algorithm for time-domain nmr data quantification. *J. Magnet. Resonance A*, **110**, 228–237, 1994.

M. Wei and G. Majda. A new theoretical approach for prony's method. *Linear Algebra and its Applications*, **136**, 119–132, 1990.