Eigenvalue Locations of Generalized Companion Predictor Matrices

Licio H. Bezerra* Fermin S. V. Bazán†

Abstract Generalized predictor companion matrices arise in the linear prediction approach for the fit of a weighted sum of n exponentials to a given set of data points. They are special solutions of matrix equations of the type $\mathsf{H}(l+p)\,\mathsf{S} = \mathsf{H}(l)$ where for each $l \geq 0$ $\mathsf{H}(l)$ is a $M \times N$ Hankel matrix obtained from this data ($M \geq N > n$). We discuss in this paper results about the eigenvalue locations of this class of solutions by means of linear algebra techniques. An application of these results in the case of all the exponents have either negative or positive real part is that the n exponentials can correspond to eigenvalues which are outside the unit circle depending on the choice of generalized predictor companion matrices. The other $(N-n)$ eigenvalues of these matrices always lie inside the unit circle and approach zero when p increases. This separation can facilitate their numerical calculation.

Key words. companion matrices, eigenvalues, linear prediction, exponential approximation

AMS subject classifications. 15A18, 65F15

1. Introduction. The identification of the parameters of functions

$$h(t) = r_1 e^{s_1 t} + \ldots + r_n e^{s_n t}, \quad Re\, s_i < 0, \quad i = 1, \ldots, n, \tag{1}$$

*Departamento de Matemática, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina 88040-900, Brasil (licio@mtm.ufsc.br). Part of this research was performed while this author was at CERFACS, Toulouse, France, and was supported by Fundação CAPES, Brasil, grant BEX 3119/95-13.

†Departamento de Matemática, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina 88040-900, Brasil (fermin@mtm.ufsc.br). This research of this author was supported by CNPq, Brasil, grant 300487/94-0 (NV).

is a problem which has been studied by several researchers in a variety of disciplines such as signal processing, mechanical vibrations, harmonic retrieval, acoustics, nuclear magnetic resonance etc. One of the approaches to the problem is the linear prediction technique, which describes the problem as a matrix equation of the type $\mathsf{H}(l+1) = \mathsf{H}(l)\,\mathsf{S}$, where $\mathsf{H}(l)$ is a $M \times N$ Hankel matrix whose i,j entries are $h_{l+i+j-2} = h((l+i+j-2)\Delta t)$ with Δt as the sampling interval [11], [12]. This equation will be called a *prediction equation* of the system if both M and N are greater than or equal to n. Observe that, if the available data is free of noise, n is the rank of $\mathsf{H}(l)$ for any l (with respect to numerical rank determination, see e.g. [2], [6], [17]). Usually we choose $M \geq N > n$ so that $\mathsf{H}(l)$ is rank deficient and thus there are an infinite number of solutions of the prediction equation, which are called *predictor matrices* of the system. An important result is that n of the eigenvalues of S are $e^{s_1 \Delta t}, ..., e^{s_n \Delta t}$, which are called *system eigenvalues.* The parameters $r_1, ..., r_n$ are calculated once these eigenvalues are known.

In practical parameter identification problems the Hankel data matrices are corrupted by noise, which is assumed to be additive: $\tilde{\mathsf{H}}(l) = \mathsf{H}(l) + \mathcal{E}_l$. Several algorithms have been proposed for solving these problems and an analysis of some available algorithms is presented in [19]. The prediction methods are based on solutions of

$$\tilde{\mathsf{H}}(l)x = \tilde{\mathsf{H}}(l+1)e_N \tag{2}$$

where e_N is the transpose of the canonical vector $(0...01)$, for example, the Minimum-Norm method [9], in which x is the minimum norm least squares solution. The different possible choices for the dimensions M, N of $\tilde{\mathsf{H}}(l)$ and the ways of solving (2) have been originated several estimation methods. For example, in [15] the Total Least Squares (TLS) approach is applied for solving the equation (2) to diminish the noise effects from the data matrices. Other methods are the Single Shift-Invariant and the Subspace Fitting methods (see again [19]), which include several algorithms like ESPRIT [14], HTLS [20], Kung's method [10], MUSIC [16] etc. Contributions for parameter estimation problems also appear in [1], [4], [7], [11], [12], [21], among others. However, we don't intend to introduce here a better performing method for identifying parameters in the practical sense. The goal of this paper is to present theoretical results about the locations of the eigenvalues of a class of solutions of the general prediction equation $\mathsf{H}(l+p) = \mathsf{H}(l)\,\mathsf{S}$, $p \neq 0$, called here generalized companion predictor matrices. These matrices naturally arise from the concept of linear prediction. Since any solution S is an $N \times N$ matrix and $N > n$, S has *extraneous* eigenvalues in addition to the system eigenvalues. One problem, therefore, is how to distinguish the system eigenvalues from the entire spectra of S, supposing the data free of noise. Our goal is to show that if the n exponentials are such that their exponents have either negative or positive real parts then there are solutions S for which the system eigenvalues have absolute value greater than 1 while the extraneous ones have modulii less than 1. We begin with $p = 1$ and show that when S is the companion matrix $\mathcal{C}(c) = [e_2\, e_3 \dots e_N\, c]$, where e_i is the ith canonical column vector and c is the minimum 2-norm least squares solution of $\mathsf{H}(l)\, x = \mathsf{H}(l+1)\, e_N$, then the extraneous eigenvalues are located inside the unit circle. Observe that the system eigenvalues corresponding to undamped signals ($Re\, s_i > 0$) should lie outside the unit circle. It is worth emphasizing that the eigenvalue locations of companion predictor matrices was studied earlier, for example by Kumaresan [9], in the context of the analysis of zeros of linear prediction-error filter polynomials for a class of deterministic signals. Here, we state results about these locations in a linear algebra context. This is done in Section 2. Also in Section 2, we show that the companion matrix $\hat{\mathcal{C}}(b) = [b\, e_1\, e_2 \dots e_{N-1}]$, which is the solution of the backward prediction equation $\mathsf{H}(l-1) = \mathsf{H}(l)\, \mathsf{S}$, with b being the minimum 2-norm least squares solution of $\mathsf{H}(l)\, x = \mathsf{H}(l-1)\, e_1$, has the extraneous eigenvalues equal to the conjugates of the extraneous eigenvalues of $\mathcal{C}(c) = [e_2\, e_3 \dots e_N\, c]$. Now, the corresponding system eigenvalues, $e^{-s_1 \Delta t}, \dots, e^{-s_n \Delta t}$, lie outside the unit circle if all the signals are damped ($Re\, s_i < 0$). Results about eigenvalues of companion matrices can also be found in [3], [5], [8], [18]. In Section 3, we introduce the concept of generalized companion predictor matrices as a class of solutions of the equation $\mathsf{H}(l+p) = \mathsf{H}(l)\, \mathsf{S}$, $p \neq 0, 1, -1$. We also present some results about the locations of the eigenvalues of these matrices. We finish this paper with numerical examples and some remarks.

2. Companion predictor matrices. Let $\mathsf{H}(l)$ be a $M \times N$ Hankel matrix whose entries are samples of $h(t)$:

$$\mathsf{H}(l) = [\vec{\mathbf{h}}_l \;\; \vec{\mathbf{h}}_{l+1} \;\; \cdots \;\; \vec{\mathbf{h}}_{l+N-1}] = \begin{bmatrix} h_l & h_{l+1} & \cdots & h_{l+N-1} \\ h_{l+1} & h_{l+2} & \cdots & h_{l+N} \\ \vdots & \vdots & & \vdots \\ h_{l+M-1} & h_{l+M} & \cdots & h_{l+M+N-2} \end{bmatrix}. \tag{3}$$

Then by (1), for all $l \geq 0$,

$$\mathsf{H}(l) = \mathsf{V}\Lambda^l \mathsf{R}\mathsf{W}^T, \tag{4}$$

where $\mathsf{V} = \mathsf{V}(\lambda_1, \ldots, \lambda_n)$ is the $M \times n$ Vandermonde matrix described by

$$\mathsf{V} = \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_n \\ \vdots & & \vdots \\ \lambda_1^{M-1} & \cdots & \lambda_n^{M-1} \end{bmatrix}, \tag{5}$$

$\Lambda = diag(\lambda_1, \ldots, \lambda_n)$, with $\lambda_j = e^{s_j \Delta t}$, $\mathsf{R} = diag(r_1, \ldots, r_n)$, and W is the submatrix of V formed by taking its N first rows. A direct consequence of this decomposition is that, for all $l \geq 0$, $rank(\mathsf{H}(l)) = n$ whenever $M \geq N \geq n$ and $\lambda_i \neq \lambda_j$, for $i \neq j$.

A predictor matrix, that is, a matrix S such that $\mathsf{H}(l+1) = \mathsf{H}(l)\mathsf{S}$, gives the new data sample $\vec{h}_{l+N}$ from the preceding N samples $\vec{h}_l, \vec{h}_{l+1}, \ldots, \vec{h}_{l+N-1}$. Observe that there can be an infinite number of matrices S satisfying this equation and that the parameters λ_i can be found from the eigenvalues of any predictor matrix according to the following relation:

$$\begin{gathered} \mathsf{H}(l+1) = \mathsf{H}(l)\mathsf{S} \\ \Updownarrow \\ \mathsf{W}^T\mathsf{S} = \Lambda \mathsf{W}^T. \end{gathered} \tag{6}$$

LEMMA 1. Let S be a solution of (6), where W is any $N \times n$ matrix and $\Lambda = diag(\lambda_1, ..., \lambda_n)$. Then $\mathcal{N}(\mathsf{W}^T)$ is an invariant subspace under S. Furthermore, if $\lambda_i \neq 0$ for all i, $\mathsf{S}\mathcal{N}(\mathsf{W}^T) = \mathcal{N}(\mathsf{W}^T)$.

Proof. $x \in \mathcal{N}(\mathsf{W}^T) \Leftrightarrow \mathsf{W}^T x = 0 \Rightarrow 0 = \Lambda \mathsf{W}^T x = \mathsf{W}^T \mathsf{S} x \Leftrightarrow \mathsf{S}x \in \mathcal{N}(\mathsf{W}^T)$ (if $\lambda_i \neq 0$ for all i, $\Rightarrow$ can be replaced by $\Leftrightarrow$ in the above chain). $\square$

LEMMA 2. Let S be a solution of (6), where W is a full rank $N \times n$ matrix ($n \leq N$) and $\Lambda = diag(\lambda_1, ..., \lambda_n)$. Let P be a $N \times (N-n)$ matrix whose columns are an orthonormal basis of $\mathcal{N}(\mathsf{W}^T)$; and let Q be a $N \times n$ matrix whose columns are an orthonormal basis of the row space of W^H, that is, $\mathcal{N}(\mathsf{W}^T)^\perp$. Then, for $x \neq 0$, $\mathsf{P}^H\mathsf{S}\mathsf{P}x = \lambda x$ if and only if $\mathsf{S}\mathsf{P}x = \lambda \mathsf{P}x$. Moreover, $\lambda(\mathsf{S}) = \lambda(\mathsf{Q}^H\mathsf{S}\mathsf{Q}) \cup \lambda(\mathsf{P}^H\mathsf{S}\mathsf{P}) = \lambda(\Lambda) \cup \lambda(\mathsf{P}^H\mathsf{S}\mathsf{P})$.

Proof. From Lemma 1, the columns of SP form a basis for $\mathcal{N}(\mathsf{W}^T)$. Then, since P is a full rank matrix, given any vector x there is a unique vector y such that $\mathsf{P}y = \mathsf{SP}x$: $y = \mathsf{P}^H\mathsf{SP}x$. Therefore, $\mathsf{P}^H\mathsf{SP}x = \lambda x \Leftrightarrow \mathsf{SP}x = \lambda \mathsf{P}x$. Now, observe that

$$\begin{pmatrix} \mathsf{Q}^H \\ \mathsf{P}^H \end{pmatrix} \mathsf{S} \begin{pmatrix} \mathsf{Q} & \mathsf{P} \end{pmatrix} = \begin{pmatrix} \mathsf{Q}^H\mathsf{SQ} & 0 \\ \mathsf{P}^H\mathsf{SQ} & \mathsf{P}^H\mathsf{SP} \end{pmatrix}.$$

Moreover, as $\overline{\mathsf{W}} = \mathsf{QZ}$, for some $n \times n$ nonsingular matrix Z, then

$$\mathsf{Q}^H\mathsf{SQ} = \mathsf{Z}^{-H}\mathsf{Z}^H\mathsf{Q}^H\mathsf{SQ} = \mathsf{Z}^{-H}\mathsf{W}^T\mathsf{SQ} =$$
$$= \mathsf{Z}^{-H}\Lambda\mathsf{W}^T\mathsf{Q} = \mathsf{Z}^{-H}\Lambda\mathsf{Z}^H\mathsf{Q}^H\mathsf{Q} = \mathsf{Z}^{-H}\Lambda\mathsf{Z}^H. \qquad \square$$

Remark 2.1. By equation (4), the rows of $\mathsf{H}(l)$ for any l are spanned by the rows of W^T. So, the matrix Q in Lemma 2 can be calculated from a QR decomposition of $\mathsf{H}(l)^H$. By Lemma 2, the characteristic polynomial of S, $p(x)$, can be written as $p(x) = (x - \lambda_1)...(x - \lambda_n)g(x)$. Since one of our goals is to identify the λ_i, $i = 1, ..., n$, then it is important to know the properties of the $N - n$ roots of $g(x)$, which will be called *extraneous roots* from now on. Observe that $g(x)$ is the characteristic polynomial of $\mathsf{P}^H\mathsf{SP}$. We will focus our attention on a special solution S of equation (6) - the companion predictor matrix:

$$\mathcal{C} = [e_2\, e_3\, \cdots\, e_N\, c] = \begin{bmatrix} 0 & 0 & \cdots & 0 & c_0 \\ 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_{N-1} \end{bmatrix}_{N\times N} \tag{7}$$

in which the column vector c is the minimum 2-norm solution of the system:

$$\mathsf{H}(l)\, c = \mathsf{H}(l+1)e_N, \tag{8}$$

or, equivalently, of the following system:

$$\mathsf{W}^T c = \Lambda \mathsf{W}^T e_N. \tag{9}$$

PROPOSITION 1. Let $\mathsf{W} = \mathsf{W}(\lambda_1, ..., \lambda_n)$ be a Vandermonde matrix of order $N \times n$, $N > n$, where $\lambda_i \neq \lambda_j$ for $i \neq j$, and let $\mathcal{C} = \mathcal{C}(c)$ be a companion matrix whose Nth column vector, $c = (c_0, ..., c_{N-1})$, is a solution of (9). If $\mu_1, ..., \mu_{N-n}$ are the extraneous roots of the characteristic polynomial of $\mathcal{C}$, then an eigenvector associated with μ_i is the vector of coefficients of the polynomial $(x-\lambda_1)...(x-\lambda_n)(x-\mu_1)...(x-\mu_{i-1})(x-\mu_{i+1})...(x-\mu_{N-n})$.

Proof. Since $\mathcal{C}$ is a solution of (6), an eigenvector associated with an extraneous root belongs to $\mathcal{N}(\mathsf{W}^T)$. Without loss of generality we may consider only the case of μ_1. So, let $a = (a_0, ..., a_{N-1})$ be an eigenvector associated with μ_1. Then $\mathcal{C}a = \mu_1 a$ is equivalent to

$$\begin{cases} c_0 a_{N-1} & = & \mu_1 a_0 \\ a_0 + c_1 a_{N-1} & = & \mu_1 a_1 \\ \vdots & \vdots & \vdots \\ a_{N-2} + c_{N-1} a_{N-1} & = & \mu_1 a_{N-1} \end{cases}$$

Since $a_{N-1} \neq 0$ (otherwise $a = 0$), we can write

$$c_0 = \frac{\mu_1 a_0}{a_{N-1}},\ c_1 = \frac{\mu_1 a_1 - a_0}{a_{N-1}}, ..., c_{N-1} = \frac{\mu_1 a_{N-1} - a_{N-2}}{a_{N-1}}.$$

On the other hand, since $a \in \mathcal{N}(\mathsf{W}^T)$,

$$a(x) = a_{N-1}x^{N-1} + ... + a_1 x + a_0 = a_{N-1}(x-\lambda_1)...(x-\lambda_n)(x-k_1)...(x-k_{N-n-1})$$

By comparing the coefficients of $a(x)$ to the coefficients of $c(x) = x^N - (c_{N-1}x^{N-1} + ... + c_1 x + c_0)$, we conclude that $c(x) = \frac{1}{a_{N-1}} a(x)(x-\mu_1)$. □

In order to prove the first proposition about the location of the eigenvalues of this matrix, we first state the following lemma, which is easily verified.

Lemma 3. Let P be a $m \times n$ matrix whose columns are orthonormal. If B is a $r \times s$ submatrix of P then $||B||_2 \leq 1$.

Proposition 2. Let $\mathsf{W} = \mathsf{W}(\lambda_1, ..., \lambda_n)$ be a Vandermonde matrix of order $N \times n$, $N > n$, where $\lambda_i \neq \lambda_j$ for $i \neq j$, and let $\mathcal{C} = \mathcal{C}(c)$ be a companion matrix whose Nth column vector, $c = (c_0, ..., c_{N-1})$, is the minimum 2-norm solution of (9). Then the $N-n$ extraneous roots of the characteristic polynomial of $\mathcal{C}$ have moduli less than 1.

Proof. Since $\mathcal{C}$ is a solution of (6), by Lemma 2 the extraneous roots are the ones of the characteristic polynomial of $\mathsf{P}^H\mathcal{C}\mathsf{P}$, where P is such that its columns form an orthonormal basis of $\mathcal{N}(\mathsf{W}^T)$. Since c is the minimum norm solution of a linear system, whose matrix is W^T, $c \in \mathcal{N}(\mathsf{W}^T)^\perp$. Hence, $\mathsf{P}^H\mathcal{C} = (\, B^H \quad 0\,)$, where B is a submatrix of P. Let μ be an extraneous root. Now, by Lemma 2 the extraneous roots are the roots of the characteristic polynomial of $\mathsf{P}^H\mathcal{C}\mathsf{P}$, and an eigenvector y of $\mathcal{C}$ associated with an extraneous eigenvalue is of the form $y = \mathsf{P}x$ for some x. If $\mu \neq 0$ then y has its last coordinate a different from zero (if not, $\mathcal{C}y = \mu y \Rightarrow y = 0$). So, if $y = \begin{pmatrix} v \\ a \end{pmatrix} = \mathsf{P}x$ is an eigenvector of $\mathcal{C}$ associated with μ such that $||y|| = 1$ (therefore $||x|| = 1$), then

$$|\mu| = |\mu x| = ||\mathsf{P}^H\mathcal{C}\mathsf{P}x|| = ||\,(\, B^H \quad 0\,)\begin{pmatrix} v \\ a \end{pmatrix}|| = ||B^H v|| \leq ||B^H||\,||v||.$$

By Lemma 3, $||B^H||_2\,||v||_2 \leq ||v||_2$. And in the case of 2-norm, $||v|| < ||\begin{pmatrix} v \\ a \end{pmatrix}|| = 1$. □

Remark 2.2. It can happen that a companion matrix as defined in Proposition 1 can have a system eigenvalue as an extraneous eigenvalue, as in the following example.

Example. Let $\lambda_1 = 1$ and λ_2 be the real root of the polynomial $2x^3 + 3x^2 + 4x + 1$. Let

$$\mathsf{W}^T = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ 1 & \lambda_2 & \lambda_2^2 \end{bmatrix}.$$

Let $\mathcal{C}(c)$ be the companion matrix such that $c = (c_0, c_1, c_2)$ is the minimum 2-norm solution of (9). Then $x^3 - c_2x^2 - c_1x - c_0 = (x - \lambda_1)(x - \lambda_2)^2$. Therefore, although λ_2 is the only extraneous root, its geometric multiplicity is 1 and not 2 because a companion matrix is nonderogatory [22], [13].

Since the parameters s_i in the function $h(t)$ have negative real part, the non-extraneous eigenvalues $e^{s_i\Delta t}$ of any solution S of (6) have moduli less than one. Our goal is to identify these eigenvalues among all the eigenvalues of S. The above proposition states that when $\mathsf{S} = \mathcal{C}(c)$ the extraneous eigenvalues also have moduli less than 1. At first sight this doesn't help us in the task of identification of the non-extraneous roots. However, when the solution $\hat{\mathcal{C}}(b)$ of the backward prediction equation

$$\mathsf{H}(l+1)\, b = \mathsf{H}(l) e_1, \tag{10}$$

is considered, the non-extraneous eigenvalues have moduli greater than 1, while the extraneous roots, which are the conjugates of the extraneous roots of $\mathcal{C}(c)$, have moduli less than one. A separation between the two sets of eigenvalues is then realized. This companion matrix, which from now on will be called companion *b-predictor* matrix, is

$$\hat{\mathcal{C}} = \hat{\mathcal{C}}(b) = [b\ e_1\ e_2\ \cdots\ e_{N-1}] = \begin{bmatrix} b_{N-1} & 1 & 0 & \cdots & 0 \\ b_{N-2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_1 & 0 & 0 & \cdots & 1 \\ b_0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{N\times N}, \tag{11}$$

which satisfies

$$\mathsf{H}(l+1)\hat{\mathcal{C}} = \mathsf{H}(l) \Leftrightarrow \mathsf{W}^T\hat{\mathcal{C}} = \Lambda^{-1}\mathsf{W}^T. \tag{12}$$

As before, the column vectors of W are left eigenvalues of $\hat{\mathcal{C}}$, which are now related to $\lambda_1^{-1}, ..., \lambda_n^{-1}$. One can see that $\hat{\mathcal{C}}(b) = \mathcal{P}^T\mathcal{C}(\hat{b})\mathcal{P}$, where $\mathcal{P} = \mathcal{P}^T$ is the $N\times N$ permutation matrix such that $\mathcal{P}e_i = e_{N-i}$ and $\hat{b} = \mathcal{P}b$.

Remark 2.3. Another result obtained in a similar way to Proposition 1 is that if $\mu_1, ..., \mu_r$ are the non-zero extraneous roots of the characteristic polynomial of $\hat{\mathcal{C}}$, the companion matrix whose first column vector, $b = (b_{N-1}, ..., b_0)$, is a solution of (12), then an eigenvector associated with μ_i is the vector of coefficients of the polynomial $(x-\lambda_1)...(x-\lambda_n)(x-\mu_1^{-1})...(x-\mu_{i-1}^{-1})(x-\mu_{i+1}^{-1})...(x-\mu_r^{-1})$. This means that if we have an extraneous root then an eigenvector associated with it gives us the other extraneous roots in addition to the system eigenvalues.

Proposition 3. Let $\mathsf{W} = \mathsf{W}(\lambda_1, ..., \lambda_n)$ be a Vandermonde matrix of order $N\times n$, $N > n$, where $\lambda_i \neq \lambda_j$ for $i \neq j$, $\lambda_i \neq 0$ for all i. Let $\hat{\mathcal{C}} = \hat{\mathcal{C}}(b)$ be a companion b-predictor matrix whose first column vector, $b = (b_{N-1}, ..., b_0)$, is the minimum 2-norm solution of (10). Then the non-extraneous eigenvalues of $\hat{\mathcal{C}}$ are λ_i^{-1}, $i = 1, ..., n$, and the $(N-n)$ extraneous roots are the conjugates of the $(N-n)$ extraneous roots of $\mathcal{C} = \mathcal{C}(c)$, the companion matrix where c is the minimum 2-norm solution of equation (9).

Proof. By (12) the column vectors of W are left eigenvectors of $\hat{\mathcal{C}}$ corresponding to the eigenvalues $\lambda_1^{-1}, ..., \lambda_n^{-1}$.

In order to see that the $N-n$ extraneous eigenvalues of $\mathcal{C}$ are the conjugates of the $N-n$ extraneous eigenvalues of $\hat{\mathcal{C}}$, let $p(z)$ and $\hat{p}(z)$ be defined as

$$p(z) = z^N - c_{N-1}z^{N-1} - ... - c_1 z - c_0$$

and

$$\hat{p}(z) = z^N - b_{N-1}z^{N-1} - ... - b_1 z - b_0,$$

respectively. Then $p(z) = f(z)g(z)$ and $\hat{p}(z) = \hat{f}(z)\hat{g}(z)$, where $f(z) = (z-\lambda_1)...(z-\lambda_n)$ and $\hat{f}(z) = (z-\lambda_1^{-1})...(z-\lambda_n^{-1})$. The problem is then to show that the roots of $g(z)$ are the conjugates of the roots of $\hat{g}(z)$. Or equivalently, that the points which minimize the function

$$F = F(z_1, ..., z_{N-n}) = 1 + |F_1(z_1, ..., z_{N-n})|^2 + ... + |F_N(z_1, ..., z_{N-n})|^2$$

are the conjugates of the points which minimize

$$\hat{F} = \hat{F}(z_1, ..., z_{N-n}) = 1 + |\hat{F}_1(z_1, ..., z_{N-n})|^2 + ... + |\hat{F}_N(z_1, ..., z_{N-n})|^2,$$

where

$$z^N - F_1 z^{N-1} - ... - F_N = f(z)(z-z_1)...(z-z_{N-n})$$

and

$$z^N - \hat{F}_1 z^{N-1} - ... - \hat{F}_N = \hat{f}(z)(z-z_1)...(z-z_{N-n}).$$

On the other hand,

$$2\pi i\, F = \int_{\mathcal{B}} \bar{z} f(z)\overline{f(z)}\, |z-z_1|^2 ... |z-z_{N-n}|^2 dz,$$

where $\mathcal{B}$ is the unit circumference. Moreover, for $z \in \mathcal{B}$, $z = \bar{z}^{-1}$, and so,

$$f(z)\overline{f(z)} = |\lambda_1|^2 ... |\lambda_n|^2 \hat{f}(\bar{z})\overline{\hat{f}(\bar{z})}.$$

But $\hat{f}(\bar{z})\overline{\hat{f}(\bar{z})} = |\bar{z}-\lambda_1^{-1}|^2 \cdots |\bar{z}-\lambda_n^{-1}|^2 = |z-\overline{\lambda_1^{-1}}|^2 \cdots |z-\overline{\lambda_n^{-1}}|^2$. Let $d(z) = (z-\overline{\lambda_1^{-1}})\cdots(z-\overline{\lambda_n^{-1}})$. So,

$$F = |\lambda_1|^2 \cdots |\lambda_n|^2 D, \text{ where } 2\pi i\, D = \int_{\mathcal{B}} \bar{z} d(z)\overline{d(z)}\, |z-z_1|^2 ... |z-z_{N-n}|^2 dz.$$

Therefore, F and D have the same minimum points. But the minimum points of D are the conjugates of the ones of $\hat{F}$. □

Remark 2.4. Since the eigenvalues of a matrix depend continuously on its entries, given a companion b-predictor matrix $\hat{\mathcal{C}}(b)$, where b is the minimum 2-norm solution of the prediction system, there is a neighbourhood $\mathcal{V}$ of b in $\mathbb{C}^N$ such that the companion matrices $\hat{\mathcal{C}}(\hat{b})$ still have n eigenvalues with moduli greater than 1 and $(N-n)$ eigenvalues with moduli less than 1 for all $\hat{b} \in \mathcal{V}$.

3. Generalized companion predictor matrices. We shall now generalize the notion of a companion predictor matrix. Since for all $l \geq 0$ we have $\mathsf{H}(l+1) = \mathsf{H}(l)\mathcal{C}$, then for any positive integer p, $\mathsf{H}(l+p) = \mathsf{H}(l+p-1)\mathcal{C} = \mathsf{H}(l+p-2)\mathcal{C}^2 = \cdots = \mathsf{H}(l)\mathcal{C}^p$. So, it is possible to compute the state $l+p$ directly from the state l. This motivates the following definition:

DEFINITION. Let $\mathsf{H}(l)$ be a Hankel matrix defined as in (3). Let p be a positive integer. S_p is a *p-predictor* matrix if, for all $l \geq 0$, $\mathsf{H}(l+p) = \mathsf{H}(l)\mathsf{S}_p$.

From the above definition and (4), we have

$$\mathsf{H}(l+p) = \mathsf{H}(l)\mathsf{S}_p \Leftrightarrow \mathsf{W}^T\mathsf{S}_p = \Lambda^p\mathsf{W}^T. \tag{13}$$

There is a collection of matrices which satisfy the above definition, and for all these matrices S_p, W^T is a matrix of left eigenvectors of S_p associated with $\lambda_1^p, \ldots, \lambda_n^p$. We shall analyze the eigenvalue location of a class of p-predictor matrices, $1 \leq p < N$, which in some sense are a sort of generalized companion matrices:

$$\mathcal{C}_p = [e_{p+1} \ \cdots \ e_N \ c^{(1)} \ \cdots \ c^{(p)}]. \tag{14}$$

Here, $c^{(i)}$, $i = 1, \ldots, p$, are column vectors satisfying the following system:

$$\mathsf{H}(l)c^{(i)} = \mathsf{H}(l+p)e_{N-i+1} \tag{15}$$

or, equivalently,

$$\mathsf{W}^T c^{(i)} = \Lambda^i \mathsf{W}^T e_N, \ i = 1, \ldots, p. \tag{16}$$

PROPOSITION 4. Let $\mathcal{C}_p$ be a p-predictor matrix defined as in (14), with the column vectors $c^{(i)}$, $1 \leq i \leq p$, being the minimum 2-norm solutions of the system (16), where $\mathsf{W} = \mathsf{W}(\lambda_1, ..., \lambda_n)$ is a $N \times n$ Vandermonde matrix with $\lambda_i \neq \lambda_j$ if $i \neq j$, and $\Lambda = diag(\lambda_1, ..., \lambda_n)$. Then $\lambda_1^p, \ldots, \lambda_n^p$ are n of its eigenvalues and the extraneous roots have moduli less than 1. Moreover, if $1 \leq p \leq n$ then $rank(\mathcal{C}_p) \leq n$; otherwise, $rank(\mathcal{C}_p) \leq N + n - p$.

Proof. Since $\mathcal{C}_p$ is also a solution of the equation $\mathsf{H}(l+p) = \mathsf{H}(l)\mathsf{S}_p$, we have by (13) that

$$\mathsf{W}^T \mathcal{C}_p = \Lambda^p \mathsf{W}^T. \tag{17}$$

Therefore, $\lambda_1^p, \ldots, \lambda_n^p$ are eigenvalues of $\mathcal{C}_p$. The proof that the extraneous roots have moduli less than 1 is analogous to the proof of Proposition 2, remarking that now the last p coordinates of the eigenvectors of $\mathcal{C}_p$ cannot be all null at the same time.

If $1 \leq p \leq n$, the set of vectors $\Lambda^i \mathsf{W}^T e_n$, $i = 1, ...p$, is linearly independent. Let $(\mathsf{W}^T)^\dagger$ be the pseudo-inverse of W^T. Then, since $(\mathsf{W}^T)^\dagger$ is a full rank matrix, the vectors $c^{(i)} = (\mathsf{W}^T)^\dagger \Lambda^i\, \mathsf{W}^T e_N$, $i = 1, \ldots, p$, are linearly independent. So, $p \leq rank\ (\mathcal{C}_p) \leq n$. For $n < p \leq N$, since for all i, $1 \leq i \leq p$, $c^{(i)} \in \mathcal{N}(\mathsf{W}^T)^\perp$, and $dim\, \mathcal{N}(\mathsf{W}^T)^\perp = n$, $rank\ (\mathcal{C}_p) \leq N + n - p$. □

When linear prediction is carried out in the backward direction we have the following definitions.

Definition. Let $\mathsf{H}(l)$ be a Hankel matrix defined as in (3). S_p^b is said a *p-backward* predictor matrix if, for $p > 0$ and $l \geq 0$, $\mathsf{H}(l) = \mathsf{H}(l+p)\mathsf{S}_p^b$.

Analogous to the case of forward prediction, if $(\forall i)\ \lambda_i \neq 0$ and $\lambda_i \neq \lambda_j$, for $i \neq j$, then

$$\mathsf{H}(l) = \mathsf{H}(l+p)\mathcal{C}_p^b \Leftrightarrow \mathsf{W}^T \mathsf{S}_p^b = \Lambda^{-p} \mathsf{W}^T. \tag{18}$$

Hence, $\lambda_1^{-p}, \ldots, \lambda_n^{-p}$ belong to the spectrum of S_p^b. Again, we are interested in analyzing p-backward predictor matrices of the type:

$$\mathcal{C}_p^b = [b^{(p)}\ \cdots\ b^{(1)}\ e_1\ \cdots\ e_{N-p}], \tag{19}$$

where b_i is the minimum norm solution of

$$\mathsf{H}(l+p) b^{(i)} = \mathsf{H}(l) e_i,\ i = 1, ..., p, \tag{20}$$

which is equivalent to the following equation:

$$\mathsf{W}^T b^{(i)} = \Lambda^{-i} \mathsf{W}^T e_1,\ i = 1, \ldots, p. \tag{21}$$

PROPOSITION 5. Let $\mathsf{W} = \mathsf{W}(\lambda_1, ..., \lambda_n)$ be a Vandermonde matrix of order $N \times n$, $N > n$, where $\lambda_i \neq \lambda_j$ for $i \neq j$, $\lambda_i \neq 0$ for all i. Let $p \leq N$ and $\mathcal{C}_p^b = [b^{(1)} \; \cdots \; b^{(p)} \; e_1 \; \cdots \; e_{N-p}]$, where $b^{(i)}$ is the minimum 2-norm solution of (21). Then, $\lambda_1^{-p}, \ldots, \lambda_n^{-p}$ are n of its eigenvalues, whereas the extraneous roots are the conjugates of those of its corresponding matrix $\mathcal{C}_p$ (equation (14)).

Proof. As $\mathcal{C}_p^b$ is also a solution of the equation $\mathsf{H}(l+p)\mathsf{S}_p^b = \mathsf{H}(l)$, we have by (18) that

$$\mathsf{W}^T \mathcal{C}_p^b = \Lambda^{-p} \mathsf{W}^T. \tag{22}$$

Therefore, $\lambda_1^{-p}, \ldots, \lambda_n^{-p}$ are eigenvalues of $\mathcal{C}_p^b$.

$\mathcal{N}(\mathsf{W}^T)$ is an invariant right subspace of both $\mathcal{C}_p$ and $\mathcal{C}_p^b$ (Lemma 1), which is associated with the extraneous roots (Lemma 2). Now, let $\mathcal{U} = (u_i^j)$ be a matrix whose $(N-n)$ columns form an orthonormal basis of $\mathcal{N}(\mathsf{W}^T)$. We wish to prove that $\mathcal{U}^H \mathcal{C}_p \mathcal{U}$ is equal to the conjugate of $\mathcal{U}^H \mathcal{C}_p^b \mathcal{U}$, which yields the statement about the extraneous roots. Since the last p columns of $\mathcal{C}_p$ as well as the first p columns of $\mathcal{C}_p^b$ are in $\mathcal{N}(\mathsf{W}^T)^\perp$, we have

$$\mathcal{U}^H \mathcal{C}_p \mathcal{U} = \begin{pmatrix} \bar{u}_{p+1}^1 & \cdots & \bar{u}_N^1 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \bar{u}_{p+1}^{N-n} & \cdots & \bar{u}_N^{N-n} & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} u_1^1 & \cdots & u_1^{N-n} \\ \vdots & & \vdots \\ u_{N-p}^1 & \cdots & u_{N-p}^{N-n} \\ u_{N-p+1}^1 & \cdots & u_{N-p+1}^{N-n} \\ \vdots & & \vdots \\ u_N^1 & \cdots & u_N^{N-n} \end{pmatrix}$$

and

$$\mathcal{U}^H \mathcal{C}_p^b \mathcal{U} = \begin{pmatrix} 0 & \cdots & 0 & \bar{u}_1^1 & \cdots & \bar{u}_{N-p}^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \bar{u}_1^{N-n} & \cdots & \bar{u}_{N-p}^{N-n} \end{pmatrix} \begin{pmatrix} u_1^1 & \cdots & u_1^{N-n} \\ \vdots & & \vdots \\ u_p^1 & \cdots & u_p^{N-n} \\ u_{p+1}^1 & \cdots & u_{p+1}^{N-n} \\ \vdots & & \vdots \\ u_N^1 & \cdots & u_N^{N-n} \end{pmatrix}.$$

So,

$$\mathcal{U}^H \mathcal{C}_p^b \mathcal{U} = (\mathcal{U}^H \mathcal{C}_p \mathcal{U})^H. \qquad \square$$

Remark 3.1. Observe that the above demonstration provides another proof of Proposition 3.

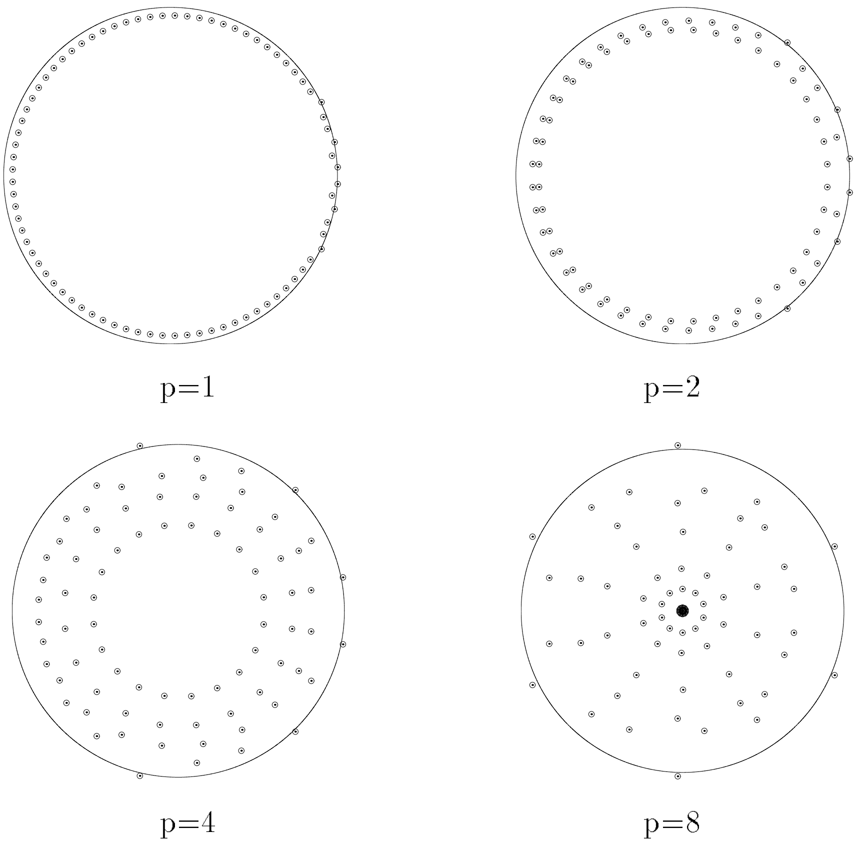


Figure 1: Eigenvalue Locations of p-Backward Predictor Matrices

Remark 3.2. When the eigenvalues λ_i are either real or complex conjugate pairs, both matrices $\mathcal{C}_p$ and $\mathcal{C}_p^b$ are real, and so, those invariant subspaces are also real.

Remark 3.3. If $\mathcal{U}$ is an orthonormal basis of $\mathcal{N}(\mathsf{W}^T)$, both $(\hat{\mathcal{C}})^p$ and $\mathcal{C}_p^b$ can be decomposed as

$$(\hat{\mathcal{C}})^p = (I - \mathcal{U}\mathcal{U}^H)\,(\hat{\mathcal{C}})^p + \mathcal{U}\mathcal{U}^H\,(\hat{\mathcal{C}})^p = \mathcal{L} + \mathcal{M} \tag{23}$$

$$\mathcal{C}_p^b = (I - \mathcal{U}\mathcal{U}^H)\,\mathcal{C}_p^b + \mathcal{U}\mathcal{U}^H\,\mathcal{C}_p^b = \mathcal{L} + \hat{\mathcal{M}} \tag{24}$$

where $\mathcal{U}\mathcal{U}^H$ is the orthogonal projection onto $\mathcal{N}(\mathsf{W}^T)$, and $I - \mathcal{U}\mathcal{U}^H$ is the orthogonal projection onto $\mathcal{N}(\mathsf{W}^T)^\perp$. Observe that $I - \mathcal{U}\mathcal{U}^H = (\mathsf{W}^T)^\dagger \mathsf{W}^T$ and $(\mathsf{W}^T)^\dagger \mathsf{W}^T\,(\hat{\mathcal{C}})^p = (\mathsf{W}^T)^\dagger \mathsf{W}^T\,\mathcal{C}_p^b = (\mathsf{W}^T)^\dagger \Delta^{-p}\,\mathsf{W}^T = \mathcal{L}$, because both

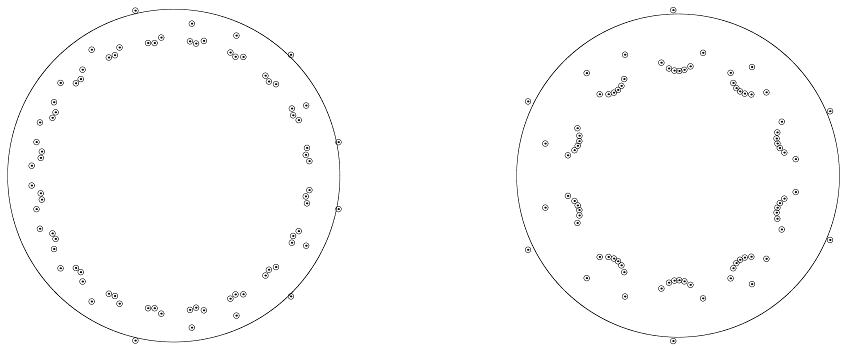


Figure 2: Eigenvalue Locations of Powers p of a 1-Backward Predictor Matrix

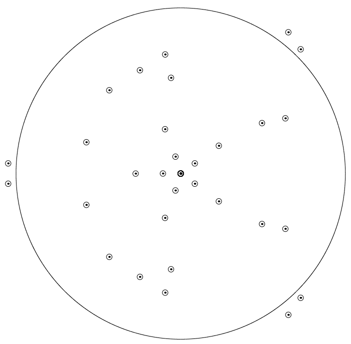
matrices are solutions of (18). Hence, $(\hat{\mathcal{C}})^p - \mathcal{C}_p^b = \mathcal{M} - \hat{\mathcal{M}}$. Since the first p columns of $\mathcal{C}_p^b$ belong to $\mathcal{N}(\mathsf{W}^T)^\perp$, the first p columns of $\hat{\mathcal{M}} = \mathcal{U}\mathcal{U}^H \mathcal{C}_p^b$ are zero. Thus, $(\hat{\mathcal{C}})^p - \mathcal{C}_p^b = \mathcal{M} - \mathsf{N}^p$, where N is the nilpotent matrix such that $\mathsf{N}e_1 = 0$, $\mathsf{N}e_i = e_{i+1}$ for $i = 1, ..., N-1$. The non-zero extraneous roots of $\mathcal{C}_p^b$ are the non-zero eigenvalues of $\mathcal{X} = \mathcal{U}\mathcal{U}^H \mathcal{C}_p^b$, which equals $\mathcal{U}\mathcal{U}^H \mathsf{N}^p$. On the other hand the non-zero eigenvalues of $\mathcal{X}$ are the non-zero eigenvalues of $\mathcal{U}\mathcal{U}^H_{(p+1:N,1:N-p)}$, that is, the submatrix of $\mathcal{U}\mathcal{U}^H$ with the rows and columns indicated by the subscript.

We hope for a better separation of the system eigenvalues from the extraneous roots of p-backward predictor matrices when p increases, because $\mathcal{U}\mathcal{U}^H \mathsf{N}^p = 0$ when $p = N$. Thus, the non-zero extraneous roots approach zero when p increases. The system signals have moduli greater than one and they will increase in magnitude with p. Observe that if $|\mu_1| \geq |\mu_2| \geq ... \geq |\mu_{N-n}|$ are the extraneous roots of $\mathcal{C}_p^b$ then

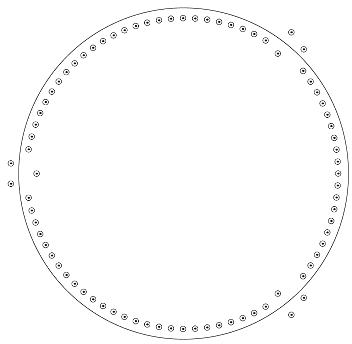
$$\sum_{i=1}^{N-n} |\mu_i|^2 \leq ||\mathcal{U}\mathcal{U}^H \mathsf{N}^p||_F^2 = \sum_{i=1}^{N-p} ||\mathcal{U}\mathcal{U}^H e_i||_2^2.$$

Since $||\mathcal{C}_p^b||_F \leq ||(\hat{\mathcal{C}})^p||_F$, we also expect a better separation between the two classes of eigenvalues when the matrix $\mathcal{C}_p^b$ is used instead of $(\hat{\mathcal{C}})^p$.

In Figure 1, we can see the behavior of the eigenvalues of $\mathcal{C}_p^b$, which is a p-backward predictor matrix for a 3-dof simulated mechanical system whose



p=16



p=1 and $\widehat{\Delta t} = 16\Delta t$

Figure 3: Eigenvalue Locations of a 16-Backward Predictor Matrix and a 1-Backward Predictor Matrix after a Downsampling Operation

impulse response function is

$$h(t) = e^{-0.06t}\sin(4t) + 0.8e^{-0.056t}\sin(t) + 1.2e^{-0.09t}\sin(9t),$$

taking $N = 80$ and $\Delta t = 0.05$. We observe in Figure 1 that the locations of the extraneous roots seem to have a certain pattern. This is because these roots are eigenvalues of submatrices $\mathcal{P}_p$ corresponding to different p values of $\mathcal{P}_0 = \mathcal{U}\mathcal{U}^H$, with each submatrix $\mathcal{P}_r$ embedded in $\mathcal{P}_s$ if $r > s$. In Figure 2 we can compare the locations of the extraneous roots of $(\hat{\mathcal{C}})^p$ with the corresponding roots of $\mathcal{C}_p^b$.

Concluding Remarks. In this paper an eigenvalue locations analysis of generalized companion predictor matrices has been presented and illustrated by numerical examples. A final remark is that if the sample rate is reduced by a factor of p ($\widehat{\Delta t} = p\Delta t$) the resulting regular predictor matrix $\widehat{\mathcal{C}}^b$ has the same system eigenvalues than the generalized companion matrix $\mathcal{C}_p^b$ (sample rate equal to Δt). However, the extraneous eigenvalues are very different. This fact is illustrated for $p = 16$ in Figure 3. The regular predictor matrix ($\widehat{\Delta t} = 0.8$) has their extraneous eigenvalues close to the unit circle, while the extraneous eigenvalues of $\mathcal{C}_{16}^b$ ($\Delta t = 0.05$) are closer to zero.

Acknowledgments. The authors wish to thank Alan McCoy (CERFACS) and Iain Duff (CERFACS/Rutherford Appleton Laboratory) for profitable discussions and helpful comments regarding the presentation of this paper.

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 Processing, 10 (1996), pp. 365-380.

[2] T. F. Chan, *Rank Revealing QR Factorizations*, Linear Algebra Appl., 88/89 (1987), pp.67-82.

[3] G. Cybenko, *Locations of Zeros of Predictor Polynomials*, IEEE Trans. on Automat. Control, AC-27 (1982), pp. 235-237.

[4] B. De Moor, *The Singular Value Decomposition and Long and Short Spaces of Noisy Matrices*, IEEE Trans. on Signal Processing, 41 (1993), pp. 2826-2838.

[5] A. Edelman and H. Murakami, *Polynomial Roots from Companion Matrix Eigenvalues*, Math. Comp., 64 (1995), pp. 763-776.

[6] G. Golub and C. F. van Loan, *Matrix Computations*, 2nd ed., The Johns Hopkins University Press, Baltimore, MD, 1989.

[7] P. De Groen and B. De Moor, *The Fit of a Sum of Exponentials to Noisy Data*, J. Comput. Appl. Math., 20 (1987), pp. 175-187.

[8] F. Kittaneh, *Singular Values of Companion Matrices and Bounds on Zeros of Polynomials*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 333-340.

[9] R. Kumaresan, *On the Zeros of the Linear Prediction-Error Filters for Deterministic Signals*, IEEE Trans. on ASSP, ASSP-31 (1983), pp. 217-221.

[10] S. Y. Kung, K. S. Arun and D. V. B. Rao, *State-Space and Singular Value Decomposition-Based Approximation Methods for the Harmonic Retrieval Problem*, J. Opt. Soc. Am., 73 (1983), pp. 1799-1811.

[11] Z. Liang and D. J. Inman, *Matrix Decomposition Methods in Experimental Modal Analysis*, J. of Vibrations and Acoustics, 112 (1990), pp. 410-413.

[12] J. Makhoul, *Linear Prediction: a Tutorial Review*, Proc. of the IEEE, 63 (1975), pp. 561-580.

[13] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Dover Publications, Mineola, NY, 1992.

[14] R. Roy and T. Kailath, *ESPRIT-Estimation of Signal Parameters Via Rotational Invariance Techniques*, IEEE Trans. Acoust. Speech. Signal Process., ASSP-37 (1989), pp. 984-995.

[15] M. A. Rahman and K. B. Yu, *Total Least Square Approach for Frequency Estimation Using Linear Prediction*, IEEE Trans. Acoust. Speech Signal Process., ASSP-35 (1987), pp. 1440-1454.

[16] R. O. Schmidt, *Multiple Emitter Location and Signal Parameter Estimation*, IEEE Trans. Antennas Propagat., AP-34 (1986), pp. 276-281.

[17] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, CA, 1990.

[18] K.-C. Toh and L. N. Trefethen, *Pseudozeros of Polynomials and Pseudospectra of Companion Matrices*, Numer. Math., 68 (1994), pp. 403-425.

[19] A. Van Der Veen, E. F. Deprettere, A. Lee Swindlehurst, *Subspace-Based Signal Analysis Using Singular Value Decomposition*, Proc. of the IEEE 81 (1993), pp. 1277-1309.

[20] S. Van Huffel, H. Chen, C. Decanniere and P. Van Hecke, *Improved Total Least Squares Based Algorithm for Time-Domain NMR Data Quantification*, J. Magn. Reson. A 110 (1994), pp.228-237.

[21] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, SIAM, Philadelphia, 1991.

[22] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, 1965.