Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Sound and Vibration 274 (2004) 433–444
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Letter to the Editor

Eigensystem realization algorithm (ERA): reformulation and system pole perturbation analysis

F.S.V. Bazán

Department of Mathematics, Federal University of Santa Catarina, 88040-900 Florianópolis SC, Brazil

Received 3 September 2002; accepted 3 September 2003

1. Introduction

The identification of modal parameters from measured data can be done in a wide variety of ways through many time and frequency domain algorithms. A list of modal parameter identification algorithms include the least-squares complex exponential algorithm [1], the polyreference time and frequency domain algorithm [2,3], Ibrahim time domain algorithm [4], and eigensystem realization algorithm (ERA) [5], among others. Presently, most of these algorithms are well understood but the effect of measurement noise on the identified parameters remains largely unexplored.

An analysis of the effect of external noise on system poles identified by ERA will be given in this work. Based on system realization and the singular value decomposition (SVD), ERA constructs a discrete state-space model of minimal order that fits measured impulse response functions (IRFs) handling closely spaced frequencies within a certain accuracy. Since its appearance in 1985, ERA has become a recognized and successful method for analyzing data in a number of engineering applications. Despite this however, very little has been done on sensitivity of system poles to noise. Some indicators of modal purity of computed parameters are given in Ref. [6] but they do not explain the pole sensitivity problem. The interest for this analysis is thus supported by the excellent reputation of ERA amongst practitioners and the lack of theoretical explanation of its robustness.

This paper provides two major contributions. The first is a detailed analysis concerning sensitivity of poles, intended to explain when and under what conditions the system poles are less sensitive to noise, as well as to provide theoretical explanation of the well observed fact that poles extracted by ERA from MIMO systems are less sensitive to noise than those extracted from a single input. The second contribution provides estimates for the pole error in the form of upper bounds. As a result, it is shown that poles near the unit circle become quite insensitive to noise whenever the dimension of the Hankel matrix is large enough and the poles themselves are not

E-mail address: fermin@mtm.ufsc.br (F.S.V. Bazán).

doi:10.1016/j.jsv.2003.09.037

extremely close to each other. Technically speaking, this result says that poles of slightly damped systems become rather insensitive to noise, under appropriate conditions.

The paper is organized as follows. Section 3 presents a reformulation of ERA which avoids the SVD, as well as some properties involving Hankel matrices and the underlying mathematical model used in modal parameter identification. The resulting reformulation is then used in Section 3 in which the main contributions of the work are described. The theoretical results of the paper are illustrated in Section 4. Finally, some conclusions are provided in Section 5.

2. Reformulating ERA

The equation of motion of a multiple-input multiple-output (MIMO) dynamic system can be represented by a set of differential equations given by

$$M\ddot{u} + C\dot{u} + Ku = f, \tag{1}$$

where M, C and K are the mass, damping and stiffness matrices respectively; u is a vector of displacement and f a vector of forcing functions. If q inputs and p outputs are available, an IRF $h(t) \in \mathbb{R}^{p\times q}$ may be found in order to describe the characteristics of the system. The following relationship between IRF's and modal parameters is known to hold [1,5,7–11]:

$$h(t) = \phi e^{St} L. \tag{2}$$

Here, $\phi \in \mathbb{C}^{p\times 2n}$ is a mode shape matrix, $S \in \mathbb{C}^{2n\times 2n}$ a diagonal matrix containing the system eigenvalues s_j, $L \in \mathbb{C}^{2n\times q}$ a modal participation factor matrix, and n the number of modes.

Let $H_{rs}(\ell)$ ($\ell \geqslant 0$) denote an $M \times N$ Hankel-block matrix with $M = r \times p$ and $N = s \times q$, whose block-entry on the position (i,j) is $h_{\ell+i+j-1} = h((\ell + i + j - 1)\Delta t)$, where Δt is the sampling interval. It is known that this Hankel-block matrix can be factorized as

$$H_{rs}(\ell) = \begin{bmatrix} \phi \\ \phi\Lambda \\ \vdots \\ \phi\Lambda^{r-1} \end{bmatrix} \Lambda^{\ell}[L\Lambda L \cdots \Lambda^{s-1}L] = \mathcal{O}_r \Lambda^{\ell} \mathcal{C}_s, \tag{3}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n, \bar{\lambda}_1, \ldots, \bar{\lambda}_n)$, $\lambda_j = e^{s_j\Delta t}$, $j = 1 : n$, and $\mathcal{O}_r$ and $\mathcal{C}_s$ are the so-called extended controllability and observability matrices, respectively. If the system is controllable and observable, $\mathcal{O}_r$ and $\mathcal{C}_s$ are both of rank $2n$ [1,9]. Thus, for $r, s \geqslant 2n$, $\mathrm{rank}(H_{rs}(\ell)) = 2n, \forall \ell \geqslant 0$. Notice that for $p = q = 1$, the above factorization reduces to

$$H_{rs}(\ell) = W_M^{\mathrm{T}} \Lambda^{\ell} R W_N, \tag{4}$$

where W_M denotes a $2n \times M$ Vandermonde matrix with entries on the jth column being given by $[\lambda_1^{j-1}, \ldots, \lambda_n^{j-1}, \bar{\lambda}_1^{j-1}, \ldots, \bar{\lambda}_n^{j-1}]^{\mathrm{T}}$, $R = \mathrm{diag}(r_1, \ldots, r_n, \bar{r}_1, \ldots, \bar{r}_n)$, Λ is as above, and W_N is the submatrix of W_M consisting of its first N columns.

Let

$$H_{rs}(\ell) = [U_1 U_2] \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{bmatrix} V_1^{\mathrm{T}} \\ V_2^{\mathrm{T}} \end{bmatrix} = U_1 \Sigma_1 V_1^{\mathrm{T}} \tag{5}$$

be a partitioned SVD of $H_{rs}(\ell)$ where $U_1 \in \mathbb{C}^{M\times 2n}$, $V_1 \in \mathbb{C}^{N\times 2n}$, with Σ_1 containing the non-zero singular values of $H_{rs}(\ell)$ in decreasing order, i.e., $\sigma_1 \geqslant \cdots \geqslant \sigma_{2n}$. Using this SVD for fixed ℓ, ERA extracts system poles (i.e., the eigenvalues $\lambda_j = \mathrm{e}^{s_j \Delta t}$, $j = 1 : 2n$) from a system matrix defined as

$$A = \Sigma_1^{-1/2} U_1^{\mathrm{T}} H_{rs}(\ell + 1) V_1 \Sigma_1^{-1/2}. \tag{6}$$

Additionally, ERA constructs an input matrix B and an output matrix C defined by

$$B = \Sigma^{1/2} V_1^{\mathrm{T}} E_q \quad \text{and} \quad C = E_p^{\mathrm{T}} U_1 \Sigma^{1/2}, \tag{7}$$

respectively. Here $E_p^T = [I_p \; 0]_{p\times M}$, where I_p is a $p \times p$ identity matrix; matrix E_q is defined analogously. The triplet $\{A, B, C\}$ is then referred to as a realization of the system with impulse response function $h(t)$ as described in Eq. (2), in the sense that $h_{k+1} = CA^k B, k \geqslant 0$.

In order to reformulate ERA, notice that $H_{rs}(\ell + 1)$ and $H_{rs}(\ell)$ are related by a matrix equation of the form

$$H_{rs}(\ell + 1) = H_{rs}(\ell) G, \tag{8}$$

where G is a block-companion matrix defined as

$$G = \begin{bmatrix} 0 & 0 & \dots & 0 & X_1 \\ I_q & 0 & \dots & 0 & X_2 \\ 0 & I_q & \dots & 0 & X_3 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \dots & I_q & X_s \end{bmatrix} \tag{9}$$

so that its last column-block has components $X_j \in \mathbb{R}^{q\times q}$ and satisfies the matrix equation

$$H_{rs}(\ell) X = b, \tag{10}$$

where b is the last column-block of $H_{rs}(\ell + 1)$. Relations (5) and (8) show that matrix A in Eq. (6) can be rewritten as

$$A = \Sigma_1^{1/2} V_1^{\mathrm{T}} G V_1 \Sigma_1^{-1/2}, \tag{11}$$

which in turn shows that system poles can be extracted from the spectrum of $V_1^{\mathrm{T}} G V_1$ as well. Matrices of this type were introduced by Bazán and Toint [12] and called predictor matrices obtained by orthogonal projection. Notice that only the product $V_1^{\mathrm{T}} X$ needs to be carried out in order to form the product $V_1^{\mathrm{T}} G$. Using Eq. (5) again, this results in

$$V_1^{\mathrm{T}} X = \Sigma_1^{-2} V_1^{\mathrm{T}} H_{rs}(\ell)^{\mathrm{T}} b. \tag{12}$$

Using Eq. (5) once more, matrix C in Eq. (7) can be rewritten as

$$C = E_p^{\mathrm{T}} H_{rs}(\ell) V_1 \Sigma_1^{-1/2}. \tag{13}$$

A simplified version of ERA algorithm is then given by matrices A, B and C described in Eqs. (11), (7) and (13), respectively. The new version allows an implementation of ERA using only the non-zero eigenvalues and corresponding eigenvectors of matrix $H_{rs}(\ell)^{\mathrm{T}} H_{rs}(\ell)$ (the right singular vectors), as opposed to the original version which requires the $2n$ triplets $\{\sigma_j, u_j, v_j\}, j = 1 : 2n$.

3. Perturbation analysis

Let $\tilde{H}_{rs}(\ell)$ denote the Hankel matrix formed from estimates $\tilde{h}_k = h_k + \varepsilon_k,\ \ k = 0, 1, \ldots,$ where ε_k stands for noise. Let $\tilde{H}_{rs}(\ell)$ have a SVD,

$$\tilde{H}_{rs}(\ell) = [\tilde{U}_1\, \tilde{U}_2]\mathrm{diag}(\tilde{\Sigma}_1, \tilde{\Sigma}_2)[\tilde{V}_1\, \tilde{V}_2]^{\mathrm{T}}, \tag{14}$$

where $\tilde{U}_1$ and $\tilde{V}_1$ contain singular vectors associated with the $2n$ largest singular values of $\tilde{H}_{rs}(\ell)$. ERA uses the dominant part of this decomposition to construct a system matrix as

$$\tilde{A} = \tilde{\Sigma}_1^{-1/2}\tilde{U}_1^{\mathrm{T}}\tilde{H}_{rs}(\ell + 1)\tilde{V}_1\tilde{\Sigma}_1^{-1/2}, \tag{15}$$

taking the eigenvalues $\tilde{\lambda}_j$ of $\tilde{A}$ as approximations for the exact eigenvalues λ_j. The goal of the section is to estimate how much the eigenvalues $\tilde{\lambda}_j$ can depart from the exact ones. More precisely, the goal is to derive pole error estimates of the form

$$|\tilde{\lambda}_j - \lambda_j| \leqslant \kappa_{j,q}\|\tilde{A} - A\|_2, \quad j = 1 : n, \tag{16}$$

which are known to hold from eigenvalue perturbation theory (see, e.g., Ref. [13, p. 323]). Here, $\kappa_{j,q}$, measures the sensitivity of λ_j to noise and it is referred to as the *condition number* of λ_j; index q is used to highlight the dependence of $\kappa_{j,q}$ on the number of system inputs. Recall that for general $A \in \mathbb{C}^{n\times n}$ with distinct eigenvalues, the condition number κ_j of eigenvalue λ_j with left eigenvector u_j and right eigenvector v_j, is defined by

$$\kappa_j = \frac{\|u_j\|_2\|v_j\|_2}{|u_j^* v_j|} \tag{17}$$

The problem with the bounds in Eq. (16) is that the error matrix $\|\tilde{A} - A\|_2$ is difficult to estimate. To overcome this drawback define $A_{\mathscr{P}}$ and $\tilde{A}_{\mathscr{P}}$ as

$$A_{\mathscr{P}} = V_1 V_1^{\mathrm{T}} G, \quad \tilde{A}_{\mathscr{P}} = \tilde{V}_1 \tilde{V}_1^{\mathrm{T}} \tilde{G}, \tag{18}$$

where V_1 and G come from Eq. (11), $\tilde{V}_1$ from the SVD of $\tilde{H}_{rs}(\ell)$ and $\tilde{G}$ is a block-companion matrix as in Eq. (9) but with its last column-block satisfying condition (12) in which exact quantities are replaced by approximations extracted from $\tilde{H}_{sr}(\ell)$. Then, using the property that discarding zero eigenvalues, the spectrum of the product of two matrices A and B satisfies $\lambda(AB) = \lambda(BA)$ (see, e.g., Ref. [13, p. 318]), it follows that the spectrum of $A_{\mathscr{P}}$ and the spectrum of $\tilde{A}_{\mathscr{P}}$ satisfy

$$\lambda(A) = \lambda(A_{\mathscr{P}}), \quad \lambda(\tilde{A}) = \lambda(\tilde{A}_{\mathscr{P}}), \tag{19}$$

and thus, the pole error obtained from $\tilde{A}_{\mathscr{P}}$ and $A_{\mathscr{P}}$ is the same as that obtained from A and $\tilde{A}$. The pole error estimates presented in the work rely on this observation and are expressed as

$$|\lambda_j - \tilde{\lambda}_j| \leqslant \kappa_{j,q}\left|\left|\tilde{A}_{\mathscr{P}} - A_{\mathscr{P}}\right|\right|_2, \quad j = 1 : n. \tag{20}$$

As it will be seen later, the main advantage of deriving pole error estimates using these inequalities instead of those in Eq. (16) is that the error matrix $\left|\left|\tilde{A}_{\mathscr{P}} - A_{\mathscr{P}}\right|\right|_2$ now is easy to estimate.

3.1. System pole sensitivity

Some results on the condition number $\kappa_{j,q}$ are given here in order to assess the sensitivity of poles related to ERA. Before this, it is worth emphasizing that $\kappa_{j,1}$ refers to a SISO system, whereas $\kappa_{j,q}$ refers to the same system but with $q>1$ inputs. Thus, all results presented below rely on the assumption that independently of the number of inputs used to realize the system, the system matrix constructed by ERA is always of the same order and always contains the same modal information. Based on this, the following properties hold.

A_1. The condition number $\kappa_{j,1}$ depends only on the Vandermonde matrix W_N and its dimension.

A_2. For $q \geqslant 1$ it holds that

$$\kappa_{j,q} = \|e_j^{\mathrm{T}} \mathscr{C}_s\|_2 \|\mathscr{C}_s^{\dagger} e_j\|_2, \quad j = 1:2n, \tag{21}$$

where $\mathscr{C}_s^{\dagger}$ is the pseudo-inverse of matrix $\mathscr{C}_s$, and e_j is the jth canonical vector in $\mathbb{R}^{2n}$.

A_3. Assume that for every $q \geqslant 1$ the number of column blocks in $\mathscr{C}_s$ is kept fixed. Then

$$1 \leqslant \kappa_{j,q} \leqslant \kappa_{j,1}, \quad j = 1:2n. \tag{22}$$

Furthermore, if $\delta_j = \min|\lambda_j - \lambda_k|, 1 \leqslant k \leqslant n, j \neq k$, then the condition number $\kappa_{j,q}$ satisfies

$$1 \leqslant \kappa_{j,q} \leqslant \left[1 + \frac{2n - 1 + \|x^{\dagger}\|_2^2 + \Pi_{j=1}^{2n}|\lambda_j|^2 - \sum_{i=1}^{2n}|\lambda_j|^2}{(2n-1)\delta_j^2}\right]^{(2n-1)/2}, \quad j = 1:2n, \tag{23}$$

where $x^{\dagger}$ is minimum 2-norm solution of the linear system (10) for $q=1$.

To prove Property A_1, observe that if $q=1$, matrix $\mathscr{C}_s$ can be rewritten as $\mathscr{C}_s = RW_N$, where $R = \mathrm{diag}(L_{1,1}, L_{2,1}, \ldots, L_{2n,1})$ is non-singular (otherwise rank $(\mathscr{C}_s) < 2n$). From this observation and (21) it follows that $\kappa_{j,1} = \|e_j^{\mathrm{T}} W_N\|_2 \|W_N^{\dagger} e_j\|_2$, which proves A_1. Property A_2 is proved in Appendix A (see Theorem A.1). Inequalities (22) and (23) are consequences of Theorem 3.2 from Ref. [14]

Properties A_1 and A_2 say that the sensitivity of system poles to noise essentially depends on the conditioning of $\mathscr{C}_s$ and that if $q=1$, the condition number $\kappa_{j,1}$ does not depend on the system input but rather on the conditioning of the Vandermonde matrix W_N introduced in Eq. (3).

Concerning property A_3, it predicts reduction in sensitivity of system poles when extracted from IRF's corresponding to multiple inputs. In practice, sensitivity of λ_j to noise reduces significantly when $q=2$ in comparison with that related to $q=1$, i.e., $q=2$ inputs usually ensures that $\kappa_{j,2} \ll \kappa_{j,1}$. This is numerically illustrated in Section 5. Inequalities (23) predict that system poles near the unit circle, but not extremely close to each other (i.e., δ_j not too small), become almost perfectly well conditioned whenever $\|x^{\dagger}\|_2^2 \approx 0$, as in this case $\kappa_{j,q} \approx 1$. The conditions $|\lambda_j| \approx 1$ and $\|x^{\dagger}\|_2^2 \approx 0$ appear frequently in connection with slightly damped systems.

A formal proof of the condition $\|x^{\dagger}\|_2^2 \approx 0$ when the dimension of the Hankel matrix is large enough is provided in Appendix A (see Theorem A.2).

3.2. Estimating the error matrix $\|A_{\mathscr{P}} - \tilde{A}_{\mathscr{P}}\|_2$

To estimate the error matrix the concept of distance between subspaces will be needed.

Definition. Let $\mathscr{S}$ and $\tilde{\mathscr{S}}$ be two subspaces in $\mathbb{R}^n$ of the same dimension, and let P and $\tilde{P}$ be orthogonal projectors onto $\mathscr{S}$ and $\tilde{\mathscr{S}}$, respectively. The distance between $\mathscr{S}$ and $\tilde{\mathscr{S}}$ is defined as

$$\mathrm{d}(\mathscr{S}, \tilde{\mathscr{S}}) = \|P - \tilde{P}\|_2 \tag{24}$$

It turns out that using the concept of subspace angles, Ref. [13, p. 603], this distance can be computed as $\mathrm{d}(\mathscr{S}, \tilde{\mathscr{S}}) = \sin(\theta)$, where θ denotes the largest canonical angle between $\mathscr{S}$ and $\tilde{\mathscr{S}}$.

Let $\mathscr{P}$ and $\tilde{\mathscr{P}}$ be the orthogonal projectors on $\mathscr{R}(V_1)$ and $\mathscr{R}(\tilde{V}_1)$, the exact and approximate controllability subspaces, respectively, i.e., $\mathscr{P} = V_1 V_1^{\mathrm{T}}$ and $\tilde{\mathscr{P}} = \tilde{V}_1 \tilde{V}_1^{\mathrm{T}}$, where V_1 and $\tilde{V}_1$ are matrices as of singular vectors as described in Eqs. (5) and (14). Let these projectors be partitioned

$$\mathscr{P} = [p_1, p_2, \ldots, p_N] \quad \text{and} \quad \tilde{\mathscr{P}} = [\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N], \tag{25}$$

where $p_i, \tilde{p}_i \in \mathbb{C}^{N\times q}, i = 1 : N$. Set $\varepsilon_i = \tilde{p}_i - p_i$ and $\eta = \check{X} - \tilde{X}$, where $\check{X}$ is the solution of minimum Frobenius norm of the linear system (10) and $\tilde{X}$ is the last column block of the block companion matrix $\tilde{G}$. Now, using the definitions of $A_{\mathscr{P}}$ and $\tilde{A}_{\mathscr{P}}$, since $A_{\mathscr{P}} - \tilde{A}_{\mathscr{P}} = [\varepsilon_2, \ldots, \varepsilon_N, \eta]$, it follows that

$$\begin{aligned}(A_{\mathscr{P}} - \tilde{A}_{\mathscr{P}})(A_{\mathscr{P}} - \tilde{A}_{\mathscr{P}})^{\mathrm{T}} &= \varepsilon_1\varepsilon_1^{\mathrm{T}} + \cdots + \varepsilon_N\varepsilon_N^{\mathrm{T}} + \eta\eta^{\mathrm{T}} - \varepsilon_1\varepsilon_1^{\mathrm{T}} \\ &= (\mathscr{P} - \tilde{\mathscr{P}})(\mathscr{P} - \tilde{\mathscr{P}})^{\mathrm{T}} + \eta\eta^{\mathrm{T}} - \varepsilon_1\varepsilon_1^{\mathrm{T}}.\end{aligned}$$

Taking 2-norm leads to the estimate

$$\|A_{\mathscr{P}} - \tilde{A}_{\mathscr{P}}\|_2^2 \leqslant \sin(\theta)^2 + \|\eta\|_2^2. \tag{26}$$

The above upper bound is important since if $\|E\|_2 = \|\tilde{H}_{rs}(\ell) - H_{rs}(\ell)\|_2 \ll \sigma_{2n}(H_{rs}(\ell))$, then $\|A_{\mathscr{P}} - \tilde{A}_{\mathscr{P}}\|_2$ becomes small, as in this case both $\sin(\theta)$ and $\|\eta\|_2$ approach 0. These conclusions rely on estimates for $\sin(\theta)$ and $\|\eta\|_2$ which depend on quantities of the form $\|E\|_2/\sigma_{2n}(H_{rs}(\ell))$ (see Ref. [15, Theorem 3.21]). Thus one is faced with the problem of investigating when the condition $\|E\|_2 \ll \sigma_{2n}(H_{rs}(\ell))$ is plausible and when is not. This is a hard matrix dimension dependent problem as both $\|E\|_2$ and $\sigma_{2n}(H_{rs}(\ell))$ monotonically increase with the dimension. The SISO case was analyzed by Bazán and Toint [18] who concluded that the desired inequality happens when the Hankel matrix is chosen so that $M \approx N$, but analyses concerning the MIMO case are lacking. Despite this, numerical experiments reported in Ref. [8] point out that if the IRFs are not dominated by the noise, the desired inequality is reached when the Hankel matrix is sufficiently large, unless severe ill-conditioning is present, i.e., unless the smallest non-zero singular value of the Hankel matrix is small.

Finally, notice that if $\|\eta\|_2^2 \leqslant \|\check{X}\|_2^2$, because this last norm approaches to zero when N is large enough, a good estimate for the error matrix is

$$\|A_{\mathscr{P}} - \tilde{A}_{\mathscr{P}}\|_2 \approx \sin(\theta). \tag{27}$$

A formal proof of the property $\|\check{X}\|_2^2 \approx 0$ for s large is given in Appendix A (see Theorem A.3).

3.3. Error estimates for system poles

In this subsection, error estimates for system poles extracted by ERA are presented and discussed. In fact, assuming $M \geqslant N \geqslant 2n$ and substituting Eqs. (23) and (26) in Eq. (20)

results in

$$|\tilde{\lambda}_j - \lambda_j| \leqslant \left[1 + \frac{2n - 1 + \|x^\dagger\|_2^2 + \Pi_{j=1}^{2n}|\lambda_j|^2 - \sum_{j=1}^{2n}|\lambda_j|^2}{(2n-1)\delta_j^2}\right]^{(2n-1)/2} \times (\sin(\theta)^2 + \|\eta\|_2^2)^{1/2}, \quad j = 1, \ldots, 2n. \tag{28}$$

Notice that both factors of the bound strongly depend on M and N. While the left factor, which measures the sensitivity of system poles to noise, requires that N be sufficiently large in order to guarantee small sensitivity; the right factor estimates the error on $A_{\mathscr{P}}$ and requires both M and N be sufficiently large to ensure that such error is small. Therefore, whenever M and N are sufficiently large, small pole errors can be expected. If, in addition to M and N being large, the poles are well separated and fall near the unit circle, then the pole error can be approximated by

$$|\tilde{\lambda}_j - \lambda_j| \approx \sin(\theta), \quad j = 1, \ldots, 2n, \tag{29}$$

unless severe noise is present. This is numerically illustrated in the next section. This conclusion results from the left factor approaching 1 and the right factor approximating $\sin(\theta)$. Poles near the unit circle appear frequently in connection with very flexible systems.

4. Numerical example: Mini-mast model

In this section the theoretical results of the paper are numerically illustrated. The system chosen, which correspond to a computer model obtained by finite element analysis of a Mini-Mast structure, is described by state equations of the form

$$\dot{x} = Ax + Bu, \quad y = cx, \tag{30}$$

where A, B and C are of orders 10×10, 10×2, and 2×10, respectively (i.e., the mathematical model considers two inputs and two outputs). Impulse response functions are thus of order 2×2 and given by

$$h_k = Ce^{A\Delta tk}B, \quad k = 0, 1, \ldots$$

For a description of the entries of the matrices A, B and C, see Ref. [16]. As described in that reference, and widely known in the spacecraft identification area, this system is sufficiently complex to be a reasonable test of pole recovering performance. The model considers five modes: two bending modes as the lowest frequency modes involving closely spaced frequencies, one torsional mode, and two additional bending modes as the highest frequency modes, involving again closely spaced frequencies. The frequencies as well as the associated damping factors (expressed as the negative real part of the eigenvalues λ_j) are given in Table 1. This table also displays the system poles $\lambda_j = e^{s_j\Delta t}$ in modulus and the separations δ_j. As in Ref. [16], the sampling rate is equal to $\Delta t = 0.03$ s.

The numerical example comprises two parts. The first is concerned with system pole sensitivity and the second with estimation of the pole error. Concerning pole sensitivity, the issues to be discussed are the influence of the number of system inputs on such sensitivity, and the behavior of the measure of sensitivity $\kappa_{j,q}$ as a function of the dimension of the controllability matrix. To this end, condition numbers $\kappa_{j,1}$ and $\kappa_{j,2}$ were computed considering controllability matrices of several

Table 1
Modal information, system poles and separations

Mode j	Damping factor	Frequency (rad/s)	$\|\lambda_j\|$	δ_j
1	0.32907	27.42011	0.99017	0.32299
2	0.38683	38.68230	0.98846	0.00982
3	0.38352	38.35103	0.98856	0.00982
4	0.09066	5.03555	0.99728	0.00011
5	0.09055	5.03176	0.99728	0.00011

Table 2
Condition numbers of system eigenvalues λ_j

Mode j	$\kappa_{j,1}$ $s=10$	$\kappa_{j,1}$ $s=20$	$\kappa_{j,2}$ $s=10$	$\kappa_{j,2}$ $s=20$
1	0.00017×10^7	0.00130×10^3	1.84786	1.00766
2	0.00127×10^7	0.02310×10^3	1.20076	1.00611
3	0.00136×10^7	0.02311×10^3	1.71432	1.00758
4	3.10889×10^7	4.75131×10^3	1.52448	1.00447
5	3.11084×10^7	4.75306×10^3	2.15234	1.00587

orders. Although the example corresponds to a MIMO system, notice that the condition numbers $\kappa_{j,1}$ can be always computed since they depend only on the Vandermonde matrix W_N rather than on the system input. These $\kappa_{j,1}$ are included here to illustrate that the sensitivity of system poles linked to ERA using multiple inputs can be much smaller than the sensitivity of the poles using a single input. Results for $s=10$ and $s=20$, which implies the Hankel matrix has $N=10$ and $N=20$ columns when $q=1$ and $N=20$ and $N=40$ when $q=2$, are displayed in Table 2 Reduction in system pole sensitivity is apparent from this table.

To illustrate the potentiality of the bounds in predicting the system pole error, square Hankel matrices $H_{rs}(0)$ of several orders were corrupted by additive zero-mean Gaussian noise, and then the estimate for the pole error given by $\sin(\theta)$, the pole error itself, $\|E\|_2$ and the smallest non-zero singular value of $H_{rs}(0)$, all were calculated from the corrupted Hankel matrix. Average values of 100 different random realizations at two noise levels are reported. The noise level was specified by the standard deviations of the random noise and was equal to 8×10^{-9} and 4.75×10^{-6}, respectively. Low noise results are displayed in Fig. 1. Fig. 1(a) shows that the smallest non-zero singular value of $H_{rs}(0)$ really increases faster than the norm of the noise $\|E\|_2$; while Fig. 1(b) illustrates the quality of $\sin(\theta)$ as an approximation for the pole error (see Eq. (29)). It is interesting to notice in this figure that the pole error itself tends to decrease as the dimension of the Hankel matrix grows.

Fig. 2 repeats Fig. 1 but with data corresponding to the high noise level obtained from Hankel matrices whose orders depend on s which ranges from $s=10$ to $s=100$. In this case, the standard deviation of the noise was calculated so as to enforce the smallest non-zero singular value of $H_{rs}(0)$ to be approximately dominated by the noise (Fig. 2(a)). Results for the pole error displayed

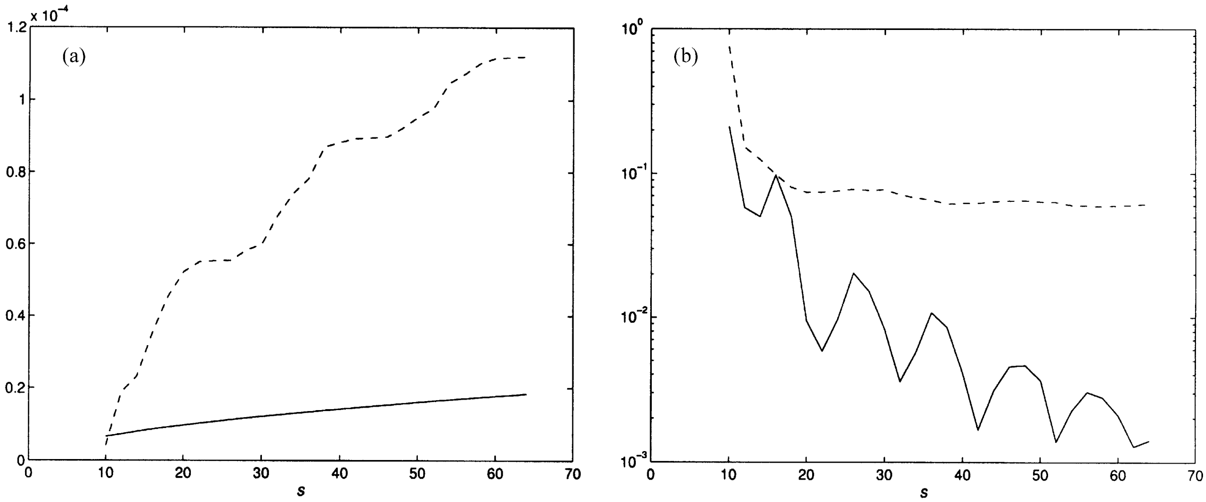

Fig. 1. Low noise results: (a) behavior of $\sigma_{2n}(H_{rs}(0))$ (dashed line) and average value of $\|E\|_2$ (solid line) for several dimensions of the Hankel matrix; (b) average value of maximum absolute error in λ_j (solid line) and average value of its estimate given by $\sin(\theta)$ (dashed line) for several dimensions of the Hankel matrix.

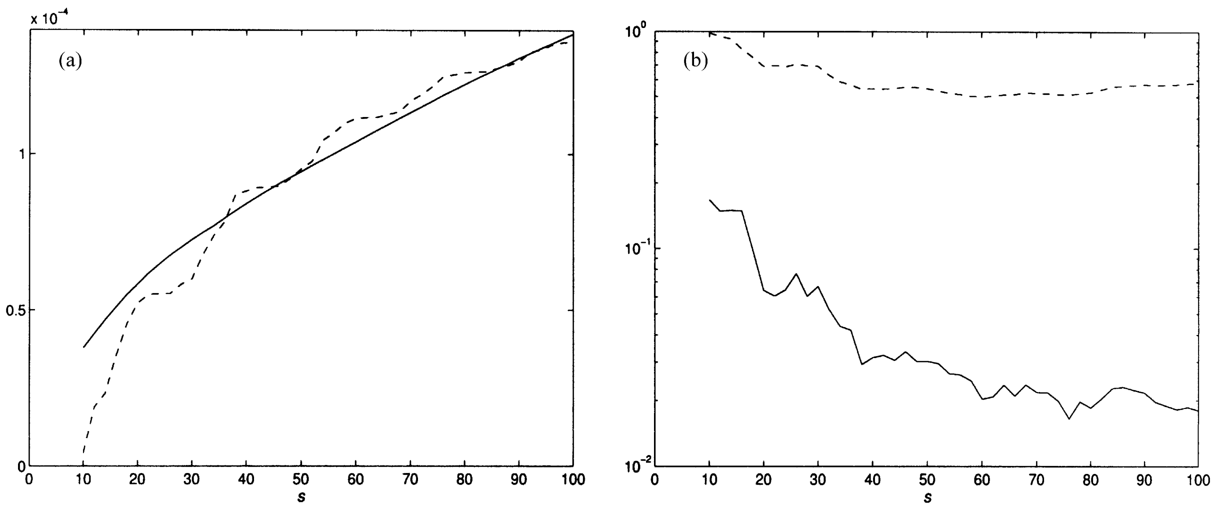

Fig. 2. High noise results: (a) behavior of $\sigma_{2n}(H_{rs}(0))$ (dashed line) and average value of $\|E\|_2$ (solid line) for several dimensions of the Hankel matrix; (b) average value of maximum absolute error in λ_j (solid line) and average values of its estimate given by $\sin(\theta)$ (dashed line) for several dimensions of the Hankel matrix.

in Fig. 2(b) show that the estimates and the pole error itself slightly deteriorate when compared to the low noise results. Finally, notice that the trend of the pole error of decreasing as a consequence of increasing the dimension of the Hankel matrix observed in the low noise situation, continues to hold even in this case.

5. Conclusions

A reformulation of the well-known eigensystem realization algorithm (ERA) was presented and an analysis on system pole sensitivity for this algorithm was carried out. The analysis relies on an existing relationship between the system matrix used by ERA and predictor matrices obtained by orthogonal projection introduced in Ref. [8,12], as well as on classical eigenvalue perturbation theory. As a result, the issue of system pole sensitivity associated with ERA was explained and estimates for the pole error in the form of upper bounds, which say much on the pole error itself, were provided. In this respect, it was concluded that poles near the unit circle become quite insensitive to noise provided the dimension of the Hankel matrix is large enough and the poles themselves are not extremely close to each other. All theoretical results were numerically illustrated using a difficult test case of pole recovering performance taken from the specialized literature.

Acknowledgements

This work was supported by $\mathrm{CNP_q}$ Brazil, grant 300487/94-O(RN).

Appendix A. Theorems

Theorem A.1. *Let $A_{\mathscr{P}}$ be as in Eq.* (8) *and let the columns of $F \in C^{N\times(N-2n)}$ form an orthonormal basis for the null subspace of $\mathscr{C}_s$. Then $A_{\mathscr{P}}$ has a spectral decomposition of the form*

$$A_{\mathscr{P}} = [\mathscr{C}_s^{\dagger} F]\begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \mathscr{C}_s \\ F^* \end{bmatrix}, \tag{A.1}$$

and therefore,

$$\kappa_{j,q} = \|e_j^{\mathrm{T}}\mathscr{C}_s\|_2 \|\mathscr{C}_s^{\dagger} e_j\|_2. \tag{A.2}$$

Proof. First notice that from the matrix equation (8) and the full-rank factorization (3) it follows that

$$\mathscr{C}_s G = \Lambda \mathscr{C}_s \tag{A.3}$$

Next, reintroduce matrix V_1 of the right singular vector of $H_{rs}(\ell)$ and observe that because $V_1 V_1^{\mathrm{T}} = \mathscr{C}_s^{\dagger}\mathscr{C}_s$, left multiplication by $\mathscr{C}_s^{\dagger}$ on both sides of Eq. (A.3) yields

$$A_{\mathscr{P}} = V_1 V_1^* G = \mathscr{C}_s^{\dagger}\mathscr{C}_s G = C_s^{\dagger}\Lambda\mathscr{C}_s. \tag{A.4}$$

From this it follows that the jth column of $\mathscr{C}^{\dagger}$ is a right eigenvector of $A_{\mathscr{P}}$ associated with eigenvalue λ_j and that the right eigenvectors associated with zero eigenvalues are vectors in $\mathscr{N}(\mathscr{C}_s)$.

Now set

$$X = [\mathscr{C}_s^{\dagger} F], \quad \text{and} \quad Y = \begin{bmatrix} \mathscr{C}_s \\ F^* \end{bmatrix}.$$

Then it follows that $XY = \mathscr{C}_s^{\dagger}\mathscr{C}_s + FF^* = \mathscr{P} + (I - \mathscr{P}) = I_N$, because FF^* is the orthogonal projector onto $\mathscr{N}(\mathscr{C}_s)$. It is also immediate that $YX = I_N$, which shows that X is a matrix of right eigenvectors of $A_{\mathscr{P}}$, thus ensuring (A.1). Equality (A.2) is an immediate consequence of Eq. (A.1) and the definition of $\kappa_{j,q}$ in Eq. (17). □

Theorem A.2. *Let $x^{\dagger}$ denote the minimum two-norm solution of the linear system*

$$H_{rs}(\ell)x = b, \tag{A.5}$$

where b is the last column of $H_{rs}(\ell + 1)$ and the Hankel matrices are formed from an IRF of a SISO system (*i.e., $p = q = 1$ in Eq.* (2)). *Then $\|x^{\dagger}\|_2$ is a decreasing function of s, and* $\lim_{s\to\infty}\|x^{\dagger}\|_2 = 0$.

Proof. Using the Vandermonde decomposition (4), it follows that

$$x^{\dagger} = H_{rs}(\ell)^{\dagger} b \quad \Leftrightarrow \quad x^{\dagger} = W_N^{\dagger}\Lambda^s e, \; e = [1, \ldots, 1]^{\mathrm{T}} \in \mathbb{R}^{2n}. \tag{A.6}$$

The assertions of the theorem follow from Eq. (A.6) upon using Theorems 2.1 and 3.8 from Ref. [17]. □

Theorem A.3. *Let $\check{X}$ denote the minimum Frobenius-norm solution of Eq.* (10) *where the Hankel matrices correspond to the MIMO case. Then $\|\check{X}\|_F \to 0$ as $s \to \infty$.*

Proof. From Eq. (10), notice that $\check{X}$ satisfices a system of linear equations with q right hand sides of type

$$[L \; \Lambda L \cdots \Lambda^{s-1}L]X = \Lambda^s L. \tag{A.7}$$

This system can be rewritten as $[L^{(1)}W_s \, L^{(2)}W_s \cdots L^{(q)}W_s]Y = \Lambda^s[L^{(1)}e \, L^{(2)} \cdots L^{(q)}e]$, where $Y = JX$, J is an appropriate permutation matrix, $L^{(i)} = \mathrm{diag}(L_{1,i}, \ldots, L_{q,i})$, $i = 1, \ldots, q$, with $L_{i,j}$ the (i,j) entry of L, and W_s as in Eq. (4). Notice that $\|Y\|_F = \|X\|_F$, since J is orthogonal. Next, consider the system

$$[L^{(1)}W_s L^{(2)}W_s \cdots L^{(q)}W_s]y = \Lambda^s L^{(i)}e, \quad i = 1, \ldots, q, \tag{A.8}$$

and notice that there exist infinitely many solutions since this is an underdetermined system. Now assume that L has no zero entry and introduce

$$\check{y}_i = \begin{bmatrix} y_1 \\ \vdots \\ y_q \end{bmatrix}, \; y_k \in \mathbb{R}^s \quad \text{such that} \quad y_k = \begin{cases} W_s^{\dagger}(L^{(i)})^{-1}Z^s L^{(i)}e, & \text{if } k = i, \\ 0 & \text{if } k \neq i, \; k = 1:q. \end{cases} \tag{A.9}$$

Thus, $\check{y}_i$ is a solution of system (A.8) and $\|\check{y}_i\|_2 = \|x^{\dagger}\|_2$. This ensures that the minimum two-norm solution of Eq. (A.8) does not exceed $\|x^{\dagger}\|_2$. Finally, if $Y = [y_1^{\dagger} \cdots y_q^{\dagger}]$, with $y_i^{\dagger}$ denoting the minimum two-norm solution of Eq. (A.8), then

$$\| Y\|_F^2 = \|y_i^\dagger\|_2^2 + \cdots + \|y_q^\dagger\|_2^2 \leqslant \|\breve{y}_i\|_2^2 + \cdots + \|\breve{y}_q\|_2^2 \leqslant q\|x^\dagger\|_2^2.$$

The assertion of the theorem is thus a consequence of Theorem A.2. If some entry $L_{i,j}$ vanishes the inverses in Eq. (A.9) can be substituted by pseudo-inverses and the proof follows in the same way. □

References

[1] R.J. Allemang, D.L. Brown, Experimental analysis and dynamic component synthesis, Vol. III—modal parameter estimation, Final Report, Department of Mechanical and Industrial Engineering, University of Cincinnati, 1987.
[2] H. Vold, T. Rocklin, The numerical implementation of a multi-input modal estimation algorithm for mini-computers *Proceedings of the International Modal Analysis Conference*, Orlando, FL, 1982, pp. 542–548.
[3] F. Lembreghts, J.L. Leuridan, H. Van Brussel, Frequency domain direct parameter identification for modal analysis: state space formulation, *Mechanical Systems and System Processing* 4 (1989) 65–76.
[4] S.R. Ibrahim, E.C. Mikulcik, A method for the direct identification of vibration parameters from the free response, *Shock and Vibration Bulletin* 47 (Part 4) (1977).
[5] J.-N. Juang, R.S. Pappa, An eigensystem realization algorithm for modal parameter identification and model reduction, *Journal of Guidance, Control and Dynamics* 8 (5) (1985) 620–627.
[6] J.-N Juang, R.S. Pappa, Effects of noise on modal parameters identified by the eigensystem realization algorithm, *Journal of Guidance* 9 (3) (1986) 294–303.
[7] R.J. Allemang, D.L. Brown, A unified matrix polynomial approach to modal parameter identification, *Journal of Sound and Vibration* 211 (3) (1998) 323–333.
[8] F.S.V. Bazán, C.A. Bavastri, An optimized pseudo inverse algorithm (OPIA) for multi-input multi-output modal parameter identification, *Mechanical Systems and Signal Processing* 10 (1996) 365–380.
[9] D.J. Ewins, *Modal Testing and Practice*, Wiley, New York, 1984.
[10] D.J. Inman, *Vibration with Control Measurements and Stability*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
[11] Q.J. Yang, et al., A system theory approach to multi-input multi-output modal parameters identification methods, *Mechanical Systems and Signal Processing* 8 (2) (1994) 159–174.
[12] F.S.V. Bazán, Ph.L. Toint, Singular value analysis of predictor matrices, *Mechanical Systems and Signal Processing* 15 (4) (2001) 667–683.
[13] G.H. Golub, C.F. Van Loan, *Matrix Computations*, 3rd Edition, The Johns Hopkins University Press, Baltimore, 1996.
[14] F.S.V. Bazán, Sensitivity analysis of prescribed eigenvalues of matrix predictor polynomials, Technical Report 12, Department of Mathematics, Federal University of Santa Catarina, Florianópolis SC, Brazil, 2002, available at http://www.mtm.ufsc.br/~cpq/pc02.html.
[15] P.C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems-Numerical Aspects of Linear Inversion*, SIAM, Philadelphia, 1998.
[16] J.-S. Lew, J.-N. Juang, R.W. Longman, Comparison of several system identification methods for flexible structures, *Journal of Sound and Vibration* 167 (3) (1993) 461–480.
[17] F.S.V. Bazán, Conditioning of Vandermonde matrices with nodes in the unit disk, *SIAM Journal of Matrix Analysis and Applications* 21 (2) (2000) 679–693.
[18] F.S.V. Bazán, Ph.L. Toint, Conditioning of infinite Hankel matrices of finite rank, *Systems and Control Letters* 41 (2000) 347–359.