

Nonnormality estimation in projection-type system realization methods

F. S. V. BAZÁN *

Department of Mathematics,
Federal University of Santa Catarina,
Florianópolis, Santa Catarina
88040-900, Brazil
Email: fermin@mtm.ufsc.br

Abstract

Using Henrici's departure from normality as tool, we investigate the influence of nonnormality on the spectral sensitivity of system matrices of projection-type system realization methods. Basically, we show that the system matrix of one of these methods is diagonally similar to the system matrix of the classical Kung's method and then conclude that the spectral sensitivity of all methods decreases significantly under condition that are often met or easy to impose in practice. Based on these results, we conclude that *all* projection-type realization methods have essentially the same spectral properties as Kung's method. Furthermore, a theoretical eigenvalue error bound is derived and illustrated by numerical simulations.

Keywords Eigenvalue sensitivity analysis, departure from normality, balanced realization, state-space models.

AMS Subject classification 65F20

1 Introduction

During the last thirty years a great deal of attention has been given to state-space models in balanced coordinates due to the role they have played in a number of important applications. Application areas include system identification, realization of systems, the construction of reduced-order models of complex systems, and sensitivity analysis [1, 2, 3, 4]. When computing reduced-order models, for example, balanced realization becomes important because it enables derivation of error bounds that inform about the quality of the model [5, 6, 7]. On the other hand, desirable properties such as system stability are often attributed to balanced state space models [8].

Increasing use of state-space models has also been observed when solving parameter estimation and harmonic retrieval problems, see, e.g., [9, 10, 11], and [12, 3, 13] for certain variants of Kung's method that have become popular in signal processing applications. Kung's method constructs

*Part of this research was performed while the author visited CERFACS, Toulouse, France, and was supported by CNPq, Brazil, grant 201407/03-5(NV).

state-space models in balanced coordinates from the singular value decomposition (SVD) of a block Hankel matrix containing noisy Markov parameters, fitting the data very well and handling clustered eigenvalues within a certain accuracy [2]. However, very little is documented on sensitivity of system eigenvalues to small perturbations in the state transition matrix.

The departure from normality of a matrix plays an important role in scientific computing: among other things, it rules the accuracy of power methods for computing eigenvalue estimates, the convergence behavior of Krylov methods, and the sensitivity of eigenvalues to matrix perturbations; concerning eigenvalue computation, for instance, it is well known that while eigenvalues of normal matrices are easy to handle, that may not be the case when computing eigenvalues of nonnormal matrices. For an exhaustive discussion on influence of nonnormality in scientific computing, the reader is referred to [14, Chap. 10].

In this paper we introduce a family of realization methods which is closely related to the classical Kung's method. Our goal is to analyze the nonnormality of the transition matrices of these methods with the purpose of deriving eigenvalue error bounds. We show that this can be done as the transition matrix of one of these methods is shown to be diagonally similar to the transition matrix of Kung's method. Our approach relies on the fact that there exists a close relationship between eigenvalue condition numbers and measures of nonnormality [14, Chap. 10]. Using this fact we show that Henrici's departure from normality of the transition matrix of the above methods can be expressed in terms of quantities which describe well the spectral sensitivity of all methods, and then we use this result for estimating the eigenvalue condition numbers.

The main conclusion of the paper is that the spectral sensitivity of *all* methods decreases significantly under conditions that are often met or easy to impose in practice. When this is the case, it is seen that the methods feature essentially the same spectral properties as Kung's method. The analysis is restricted to the single-input single-output case with the assumption that all eigenvalues are distinct.

The paper is organized as follows. Section 2 presents the proposed family of projection-type realization methods along with a brief review of Kung's method. A set of relationships and properties of system theory as well as the notation used throughout is also included in this section. In Section 3, the connection between a particular projection method and Kung's method is made precise and the departure from normality of the system matrix of all methods is analyzed. These results are then exploited to derive an eigenvalue error bound which is described and discussed in Section 4. Numerical simulations that illustrate the theory are reported in Section 5. Finally, conclusions are provided in Section 6.

2 Projection-type realization methods and review of Kung's method

In the remaining of the paper the entries of a matrix A are denoted by $a_{i,j}$ or $[A]_{i,j}$. A^* denotes the conjugate transpose of A and $A^\dagger$ its Moore-Penrose pseudo-inverse. The column (resp. row) subspace of A is denoted by $\mathcal{R}(A)$ (resp. $\mathcal{R}(A^*)$). In addition, as usual, $\|A\|_2$ and $\|A\|_F$ denote the spectral norm and the Frobenius norm of A , respectively. Finally, e_j denotes the j -th unit canonical vector of appropriate dimension; further notation will be introduced when needed.

In order to describe the methods we are interested in, we first describe some definitions and relationships from system theory. Consider a linear, discrete, time invariant system with p inputs

and q outputs described in state-space form as:

$$\begin{cases} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k + Du_k \end{cases}, \quad (1)$$

where the system matrix (or state transition matrix) A is $n \times n$, the control matrix B is $n \times p$, C is $q \times n$ and D is $q \times p$. We always assume that the system is stable, i.e., the eigenvalues λ_j of A fall inside the unit circle, and for simplicity, that D is the zero matrix. The estimation of A , B , and C (up to a similarity transformation) from either input/output data $\{u_k, y_k\}$ or Markov parameters h_k :

$$h_k = CA^{k-1}B, \quad k = 1, \dots, h_L,$$

commonly known as system identification or realization problem, is of paramount importance in system theory. The proposed system realization methods, to be described in the sequel, rely on the use of Markov parameters as data and start by forming an $(M \times q) \times (N \times p)$ block Hankel matrix $H(\ell)$ with block entries in the position (i, j) equal to $h_{i+j+\ell-2}$. Recall that for controllable and observable systems and $M, N \geq n$, this matrix has a full-rank factorization of type

$$H(\ell) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{M-1} \end{bmatrix} A^{(\ell-1)} [B \ AB \ AB^2 \ \dots \ A^{N-1}B] = \mathcal{O}A^{(\ell-1)}\mathcal{C}, \quad \ell \geq 1 \quad (2)$$

where $\mathcal{C}$ and $\mathcal{O}$ are the so-called extended controllability and extended observability matrices, respectively, and that for single-input single-output systems the above factorization becomes

$$H(\ell) = V_M R \Lambda^{\ell-1} W_N, \quad (3)$$

where

$$W_N = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{N-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{N-1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \dots & \lambda_n^{N-1} \end{bmatrix}, \quad (4)$$

$$R = \text{diag}(r_1, r_2, \dots, r_n), \quad \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n), \quad \text{and} \quad V_M = W_M^T.$$

We now observe that if $G \in \mathbb{R}^{(M \times q) \times (M \times q)}$ is any solution of the matrix equation

$$H(2) = GH(1), \quad (5)$$

from (2) it is straightforward to see that $\lambda(G)$, the spectrum of G , is described by

$$\lambda(G) = \{\lambda_1, \dots, \lambda_n\} \bigcup \{\lambda_k^{(s)}\}_{k=1}^{M \times q - n}$$

where the λ_j 's are the system eigenvalues and $\lambda_k^{(s)}$ are spurious eigenvalues which depend on the solution chosen. The main difficulty encountered in connection with the use of G for purpose of

system identification is the link of the system eigenvalues with a “large eigenvalue problem” and the separation of them from the entire spectrum of G . We shall show that this can be circumvented by using an $n \times n$ projected version of G defined by

$$\check{A} = U^*GU, \quad (6)$$

where U is any orthonormal basis of $\mathcal{R}(H(1))$. The motivation for this definition relies on the fact that the spectrum of $\check{A}$ only contains the system eigenvalues, i.e., we have that $\lambda(\check{A}) = \{\lambda_1, \dots, \lambda_n\}$. Since U can be chosen in infinitely many ways, the above result suggests that we can construct infinitely many realizations with system matrices as defined in (6). From here on, any realization based on (6) will be called a *projection-type realization*. Before continuing we emphasize that because our interest is to analyze the spectral sensitivity of the system matrix of the methods rather than the estimation of parameters B and C involved in the realization process, we will focus only on the construction and analysis of the corresponding system matrix. We will see that interesting spectral properties of matrices $\check{A}$ can be assured provided that the solution G of (5) is completely determined. This is the subject of the following proposition, the statement of which is one of the most important contributions of this work.

Proposition 1 *Assume that $M, N \geq n$ and let G be any solution of the matrix equation (5). Then, independently of the solution chosen, matrix $\check{A}$ in (6) can be rewritten as*

$$\check{A} = U^*FU, \quad (7)$$

where F is a block companion matrix with last row block defined by $X = CA^M\mathcal{O}^\dagger$ and satisfying $H(2) = FH(1)$.

Proof: For simplicity, we shall prove the statement of the proposition for the single-input single-output case; the proof for the general case is similar. Since $H(1)$ is rank deficient, any solution of the matrix equation (5) is of the form

$$G = H(2)H(1)^\dagger + Z \quad (8)$$

with Z satisfying $ZH(1) = 0$. In addition, since the columns of U span $\mathcal{R}(H(1))$, from (8) it follows that

$$\begin{aligned} U^*GU &= U^*H(2)H(1)^\dagger U \\ &= U^*V_M\Lambda V_M^\dagger U \end{aligned} \quad (9)$$

where we have used the fact that (3) is a full-rank factorization of $H(\ell)$ together with pseudo inverse properties. Now observe that

$$V_M\Lambda V_M^\dagger = \begin{bmatrix} p_2^T \\ \vdots \\ p_M^T \\ e^T \Lambda^M V_M^\dagger \end{bmatrix} \quad (10)$$

where e is a vector of all ones of appropriate dimension and p_i is the i -th column of $\mathcal{P} = V_M V_M^\dagger$, the orthogonal projector onto $\mathcal{R}(H(1))$. Let $p_i^T = e_i^T - q_i^T$, where q_i denotes the i -th column of

the orthogonal projector onto $[\mathcal{R}(H(1))]^\perp$. Substitution of (10) in (9) with p_i as before then leads to $U^*V_M\Lambda V_M^\dagger U = U^*FU$ where F is a companion matrix as described in the statement of the proposition, as claimed. ■

We end this section by reviewing Kung's method. Let

$$H(1) = \mathcal{U}\Sigma\mathcal{V}^* = [\mathcal{U}_1 \ \mathcal{U}_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathcal{V}_1^* \\ \mathcal{V}_2^* \end{bmatrix} \quad (11)$$

be a partitioned SVD of $H(1)$ where both $\mathcal{U}_1$ and $\mathcal{V}_1$ have n columns and $\Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ in which the nonzero singular values σ_i of $H(1)$ are ordered in decreasing form. Kung's method constructs a realization by using estimates of an observable (resp. controllable) matrix obtained from the singular value decomposition of matrix $H(1)$ rewritten as

$$H(1) = (\mathcal{U}_1 \Sigma_1^{1/2})(\Sigma_1^{1/2} \mathcal{V}_1^*),$$

from which the system matrix is defined by

$$A_K = \Sigma^{-1/2} \overline{U}^\dagger \underline{U} \Sigma^{1/2} = \overline{U}^\dagger \underline{U}, \quad (12)$$

where $\overline{U}$ and $\underline{U}$ (resp. $\overline{\mathcal{U}}$ and $\underline{\mathcal{U}}$) are formed by taking respectively the first N block rows and the last N block rows of $\mathcal{U}_1$ (resp. of $\mathcal{U}_1 \Sigma_1^{1/2}$).

The realization with system matrix A_K assures that the associated controllable and observable Gramians are equal to Σ_1 , in which case it is said that the system is balanced or that the state-space model is in balanced coordinates [4]. Observe that a state transition matrix can also be determined as

$$A_K = \Sigma_1^{1/2} \underline{V} \overline{V}^\dagger \Sigma_1^{-1/2}, \quad (13)$$

where $\underline{V}$ and $\overline{V}$ are obtained from $\mathcal{V}_1^*$ by taking its first $(N-1)$ block columns and its last $(N-1)$ blocks columns, respectively.

3 Nonnormality Estimation

Our goal here is to estimate Henrici's departure from normality of the system matrices of both the projection-type methods and Kung's method, focusing on single-input single-output systems with the assumption that the eigenvalues of the system are distinct. The following technical result will be needed.

Lemma 2 *Let $\check{A}$ be as in Prop. 1. Then $\check{A} = \overline{U}^\dagger \underline{U}$ where $\overline{U}$ (resp. $\underline{U}$) is formed by taking the first (resp. last) $N-1$ rows of U .*

Proof: From the full-rank matrix factorization (3) and the matrix equation $H(2) = FH(1)$ it follows that $FV_N = V_N\Lambda$. But since $\mathcal{R}(U) = \mathcal{R}(V_N)$, it follows that U is invariant under F and therefore

$$U\check{A} = UU^*FU = \mathcal{P}FU = FU,$$

where the last equation is because $\mathcal{P} = UU^*$ is the orthogonal projector onto $\mathcal{R}(H(1))$ and the columns of FU belong to such subspace. Equating the first $N - 1$ rows of the matrix equation $U\check{A} = FU$ and then solving for $\check{A}$ leads to $\check{A} = \overline{U}^\dagger \underline{U}$, as desired. $\blacksquare$

Observe that if in the previous lemma the orthonormal basis used in the definition of matrix $\check{A}$ is chosen to be $\mathcal{U}_1$, the basis obtained from the SVD of $H(1)$, then the corresponding matrix $\check{A}$ becomes diagonally similar to A_K , the system matrix of Kung's method. This observation is exploited in the following proposition to determine the Frobenius norm of matrix A_K from the entries of the system matrix of a well-chosen projection-type method.

Proposition 3 *Let $\check{A} = \mathcal{U}_1^* F \mathcal{U}_1$ where F is as before and $\mathcal{U}_1$ is the basis of $\mathcal{R}(H(1))$ obtained from the SVD of $H(1)$. Then, provided $M = N > n$ we have*

$$\|A_K\|_F^2 = \sum_i |a_{i,i}|^2 + 2 \sum_{i=1}^{n-1} \sum_{j>i} |a_{i,j}| |a_{j,i}|, \quad (14)$$

where $a_{i,j}$ are entries of $\check{A}$, and A_K is the state transition matrix of Kung's method.

Proof: Since for $M = N$ the Hankel matrix $H(1)$ is symmetric, its singular values satisfy $\sigma_j = |\gamma_j|$, $j = 1 : n$, where the γ_j are the non zero eigenvalues of $H(1)$. This implies that there exists a diagonal matrix J with entries $[J]_{j,j}$ equal to 1 or -1 such that $\mathcal{U}_1 = \mathcal{V}_1 J$. Using this relation and the definition of A_K (see (12)), because $\overline{U}^\dagger \underline{U} = \check{A} = \mathcal{U}_1^* F \mathcal{U}_1$ by Lemma 1, we obtain

$$A_K = \Sigma_1^{-1/2} J \mathcal{V}_1^* F \mathcal{V}_1 J \Sigma_1^{1/2} = J (\Sigma_1^{1/2} \mathcal{V}_1^* F^* \mathcal{V}_1 \Sigma_1^{-1/2})^* J = J \check{A}_K^* J,$$

where the last equality is for $(\mathcal{V}_1^* F^* \mathcal{V}_1)$ is the projected version of matrix $\underline{V} \overline{V}^\dagger$ in (13). As the above result shows that $\check{A}$ is J -symmetric, we have that

$$|[A_K]_{i,j}| = |[A_K]_{j,i}|, \quad 1 \leq i, j \leq n. \quad (15)$$

From this equality and the definition of A_K , it follows that

$$0 = \left(\frac{\sigma_j^{1/2}}{\sigma_i^{1/2}} |a_{i,j}| - \frac{\sigma_i^{1/2}}{\sigma_j^{1/2}} |a_{j,i}| \right)^2 = \frac{\sigma_j}{\sigma_i} |a_{i,j}|^2 + \frac{\sigma_i}{\sigma_j} |a_{j,i}|^2 - 2 |a_{i,j}| |a_{j,i}|, \quad (16)$$

where the $a_{i,j}$'s are entries of matrix $\check{A}$, and hence

$$|[A_K]_{i,j}|^2 + |[A_K]_{j,i}|^2 = 2 |a_{i,j}| |a_{j,i}|. \quad (17)$$

If the Frobenius norm of A_K is computed as

$$\|A_K\|_F^2 = \sum_{i=1}^n |[A_K]_{i,i}|^2 + \sum_{i=1}^{n-1} \sum_{j>i}^n |[A_K]_{i,j}|^2 + |[A_K]_{j,i}|^2,$$

the equality (14) is a consequence of (17). This concludes the proof. $\blacksquare$

Matrix $\check{A}$ described above via pseudo-inversion is often used in estimating parameters from exponentially-damped noisy signals [11]. This matrix is sometimes referred as to the transition matrix of Kung's method in non balanced coordinates.

An immediate consequence of (15) is that A_K is balanced in the 2-norm. Recall that a matrix is balanced in the α -norm if for any i , the α -norm of the i -th row equals the α -norm of the i -th column. Using results of Osborne [15], it follows that $\Sigma_1^{1/2}$ minimizes the Frobenius norm of all matrices $G = D^{-1}AD$ where D is a diagonal matrix, and so $\|A_K\|_F \leq \|\check{A}\|_F$. That is, the similarity transformation $G = D^{-1}AD$, with $D = \Sigma_1^{1/2}$, not only balances matrix $\check{A}$ but also reduces the Frobenius norm. As a result the nonnormality A_K and all system matrices of projection-type methods as described in Lemma 1 can be now described. This is given in the following corollary where we use as measure of nonnormality Henrici's departure from normality [16, p. 314]. Recall also that for general $A \in \mathbb{C}^{n \times n}$, Henrici's departure from normality, denoted by $\mathcal{D}(A)$, is defined as [16]

$$\mathcal{D}(A) = \|A\|_F^2 - \sum_{j=1}^n |\lambda_j(A)|^2. \quad (18)$$

Corollary 4 *Assume that $M = N > n$. Then for any projection-type realization method with system matrix A it holds*

$$\mathcal{D}(A_K)^2 \leq \mathcal{D}(A)^2. \quad (19)$$

Proof: We first prove that the inequality holds for the case where we consider the system matrix $\check{A}$. In fact, a simple expression for $\mathcal{D}(A_K)^2$ can be given using the identity

$$|a_{i,j}|^2 + |a_{j,i}|^2 = 2|a_{i,j}||a_{j,i}| + (|a_{i,j}| - |a_{j,i}|)^2$$

in (14). Doing this, (14) takes the form

$$\mathcal{D}(A_K)^2 = \mathcal{D}(\check{A})^2 - \sum_{i=1, n}^n \sum_{j>i}^n (|a_{i,j}| - |a_{j,i}|)^2. \quad (20)$$

which proves (19) when the system matrix is $\check{A}$. We now make the crucial observation that all system matrices of projection-type realization methods, as defined in (6), are unitarily similar to $\check{A}$. The assertion of the corollary then follows from this observation because all these system matrices share the same Henrici's departure from normality. $\blacksquare$

Two conclusions can be drawn from equality (20). The first is that $\mathcal{D}(A_K)^2$ is always smaller than $\mathcal{D}(\check{A})^2$ and that matrix A_K can become much more close to being normal than matrix $\check{A}$ provided that $|a_{i,j}|$ and $|a_{j,i}|$ differ by a large quantity. In terms of spectral sensitivity, this suggests that the eigenvalue problem related to A_K should be better conditioned than the eigenvalue problem related to A , despite the fact that both matrices share the same spectrum. We will return to this point later. The second conclusion is that the departure from normality of $\check{A}$ and A_K can become small

provided that $|a_{i,j}| \approx |a_{j,i}|$. This is important because in such case, both matrices $\check{A}$ and A_K are ensured to behave almost as normal matrices in the sense that their eigenvalues become relatively insensitive to small matrix perturbations. In other words, the condition $|a_{i,j}| \approx |a_{j,i}|$ ensures that $\check{A}$ and A_K essentially share the same spectral properties, the same comment being valid if $\check{A}$ is replaced by any system matrix of any projection method as those described before. The following corollary provides an estimate for $\mathcal{D}(A_K)^2$.

Corollary 5 *Using the same assumptions of Corollary 4 it holds that*

$$\mathcal{D}(A_K)^2 \leq n - 1 + \|f^\dagger\|_2^2 + \prod_{i=1}^n |\lambda_i|^2 - \sum_{i=1}^n |\lambda_i|^2 \quad (21)$$

where $f^\dagger$ denotes the minimum 2-norm solution of the linear system $H(1)f = H(2)e_N$.

Proof: We first observe that using Theorem 3 in [17] it follows

$$\mathcal{D}(A)^2 = n + \|f^\dagger\|_2^2 - \|p_1\|_2^2 - \sum_{i=1}^n |\lambda_i|^2,$$

where p_1 is the first column vector of the orthogonal projection onto $\mathcal{R}(V_N)$. The statement of the corollary is thus a consequence of Corollary 4 and the fact that the right side of (21) is an upper bound for $\mathcal{D}(A)^2$, as seen from Lemma 7 in [17]. $\blacksquare$

In reference [17] it was proved that $\|f^\dagger\|_2$ decreases with the dimension M of the Hankel matrix and that this norm approaches zero as M is going to infinite. Therefore, since the right hand of (21) is an upper bound for $\mathcal{D}(A)^2$, A being any system matrix of any projection-type realization method, we can conclude that small values of Henrici's departure from normality of *all* these matrices can be expected provided that M is large enough and the eigenvalues satisfy $|\lambda_j| \approx 1$. We end the section with the observation that the conditions $\|f^\dagger\|_2 \approx 0$ and $|\lambda_j| \approx 1$ are often met in connection with slightly damped systems [12, 18].

4 Spectral Properties

Before proceeding, recall that for general $A \in \mathbb{C}^{n \times n}$ having simple eigenvalues $\lambda_j(A)$, the *condition number* of eigenvalue $\lambda_j(A)$ is given as

$$\kappa_j = \frac{\|u_i\|_2 \|v_i\|_2}{|u_i^* v_i|}, \quad (22)$$

where v_j and u_j are right and left eigenvectors of A associated with λ_j . It is well-known that $u_i^* v_i$ is always nonzero and that the real number κ_j controls the sensitivity of eigenvalue λ_j to perturbations on A (see, e.g., Wilkinson [19, p. 87]). The following result will be needed to prove our main result concerning the sensitivity spectral of the system matrices of all realization methods studied here.

Lemma 6 *The state transition matrix of Kung's method in non balanced coordinates $\check{A}$, has a spectral decomposition given as*

$$\check{A} = (\mathcal{U}_1^* V_N) \Lambda (\mathcal{U}_1^* V_N)^{-1} = (\mathcal{U}_1^* V_N) \Lambda (V_N^\dagger \mathcal{U}_1). \quad (23)$$

Consequently, the condition number $\check{\kappa}_j$ of eigenvalue λ_j satisfies

$$\check{\kappa}_j = \|V_N e_j\|_2 \|V_N^\dagger e_j\|_2. \quad (24)$$

Proof: It is sufficient to observe that the companion matrix F used in the definition of $\check{A}$ satisfies the matrix equation $FV_N = V_N \Lambda$, and that the columns of $\mathcal{U}_1$ form an orthonormal basis for $\mathcal{R}(V_N)$. Details can be encountered in Bazán [17]. $\blacksquare$

Proposition 7 *Let $\hat{\kappa}_j$ and $\check{\kappa}_j$ and denote the condition numbers of eigenvalue λ_j regarded as eigenvalue of A_K and $\check{A}$, respectively. Then it holds that*

$$\hat{\kappa}_j \leq \check{\kappa}_j, \quad j = 1, \dots, n.$$

Proof: Let $\hat{v}_j$ and $\hat{u}_j$ denote right and left eigenvectors of matrix A_K associated with λ_j . From the definition of A_K and (23) it follows that they can be taken as

$$\hat{v}_j = \Sigma_1^{-1/2} \mathcal{U}_1^* V_N e_j, \quad \hat{u}_j^* = e_j^* V_N^\dagger \mathcal{U}_1 \Sigma_1^{1/2}. \quad (25)$$

These eigenvectors satisfy the normalization condition $\hat{u}_j^* \hat{v}_j = 1$, so $\hat{\kappa}_j = \|\hat{u}_j\|_2 \|\hat{v}_j\|_2$. From this observation and (25) it follows that using (26)

$$\begin{aligned} \|\hat{v}_j\|_2^2 &= e_j^* V_N^* \mathcal{U}_1 \Sigma_1^{-1} \mathcal{U}_1^* V_N e_j, \\ &= e_j^* V_N^* \mathcal{U}_1 \mathcal{V}_1^* \mathcal{V}_1 \Sigma_1^{-1} \mathcal{U}_1^* V_N e_j \\ &= e_j^* V_N^* \mathcal{U}_1 \mathcal{V}_1^* H(1)^\dagger V_N e_j \\ &= e_j^* V_N^* \mathcal{U}_1 \mathcal{V}_1^* W_N^\dagger R^\dagger V_N^\dagger V_N e_j \\ &= (e_j^* V_N^* \mathcal{U}_1) (\mathcal{V}_1^* W_N^\dagger e_j) / r_j, \end{aligned} \quad (26)$$

where $r_j = [R]_{j,j}$ and the last four equalities are for $\mathcal{V}_1 \Sigma_1^{-1} \mathcal{U}_1^* = H(1)^\dagger = W_N^\dagger R^\dagger V_N^\dagger$ and $V_N^\dagger V_N = I$. Similarly, from $\|\hat{u}_j\|_2$ it follows

$$\begin{aligned} \|\hat{u}_j\|_2^2 &= e_j^* V_N^\dagger \mathcal{U}_1 \Sigma_1 \mathcal{U}_1^* V_N^* e_j \\ &= e_j^* V_N^\dagger \mathcal{U}_1 \Sigma_1 \mathcal{V}_1^* \mathcal{V}_1 \mathcal{U}_1^* V_N^* e_j \\ &= e_j^* V_N^\dagger V_N R W_N \mathcal{V}_1 \mathcal{U}_1^* V_N^* e_j \\ &= r_j (e_j^* W_N \mathcal{V}_1) (\mathcal{U}_1^* V_N^* e_j), \end{aligned} \quad (27)$$

where, as before, we have used the fact that $\mathcal{U}_1 \Sigma_1 \mathcal{V}_1^* = H(1) = V_N R W_N$. From (26),(27), and the fact that $\|V_N e_j\|_2 = \|W_N e_j\|_2$ and $\|W_N^\dagger e_j\|_2 = \|V_N^\dagger e_j\|_2$, using the Cauchy-Schwartz inequality it then follows that

$$\hat{\kappa}_j^2 \leq \check{\kappa}_j^2,$$

and the proof concludes. ■

An analysis of the condition numbers $\check{\kappa}_j$ and their use to estimate the condition number of $H(1)$ can be encountered in [20, 18]. The main conclusion drawn there is that $\check{\kappa}_j$ depends on eigenvalue separation, the closeness of the eigenvalues to the unit circle, and mainly on the dimension of the Hankel matrix in that moderate condition numbers are generally obtained whenever the Hankel matrix is large enough. Following the line of analysis in [18] an upper bound for $\hat{\kappa}_j$ can be immediately obtained.

Corollary 8 *For all $N > n$, the condition number $\hat{\kappa}_j$ satisfies*

$$1 \leq \hat{\kappa}_j < \left[1 + \frac{n-1 + \|f^\dagger\|_2^2 \|p_1\|_2^2 + \prod_{i=1}^n |\lambda_i|^2 - |f_0|^2 - \sum_{i=1}^n |\lambda_i|^2}{(n-1)\delta_j^2} \right]^{(n-1)/2}, \quad j = 1 : n, \quad (28)$$

where $f^\dagger$ denotes the minimum 2-norm solution of the linear system $H(1)f = H(2)e_N$, f_0 is the first component of $f^\dagger$, p_1 denotes the first column vector of the orthogonal projection onto $\mathcal{R}(V_N)$, and $\delta_j = \min_{\substack{j \\ j \neq i}} |\lambda_i - \lambda_j|$.

Proof: From [21]

$$\hat{\kappa}_j \leq \left[1 + \frac{\mathcal{D}(A_K)^2}{(n-1)\delta_j^2} \right]^{(n-1)/2}. \quad (29)$$

Inequality (28) follows from using the expression for $\mathcal{D}(A_K)$ described in (20) where we use the exact value of $D(\check{A})^2$, as described in Proposition 3 in reference [18]. ■

Observe that the right hand of inequality (28) can grossly overestimate the condition number $\hat{\kappa}_j$ since, as commented above, $\mathcal{D}(A_K)^2$ can become much smaller than $\mathcal{D}(A)^2$ when $|a_{i,j}|$ differs largely from $|a_{j,i}|$, a fact frequently seen when M is not large enough. Numerical examples illustrating this phenomenon are presented later. Observe also that if N is large enough and the eigenvalues in modulus are reasonably close to 1 but not extremely close to each other, which is very often seen to happen in practice, then $\hat{\kappa}_j$ approaches 1 because under these conditions $\|f\|_2^2 \approx 0$. That is, if these conditions are satisfied, then the eigenvalues become relatively insensitive to small perturbation in the system matrix.

Remark A comment regarding the accuracy of Kung's method in the noisy data case is required. First, it should be noted that because A_K is obtained from $\check{A}$ via a diagonal similarity, the same set of eigenvalue are to be obtained from both matrices, at least theoretically, this result being valid regardless of whether the data are corrupted by noise or not. Secondly, it is convenient to invoke an eigenvalue perturbation analysis recently reported in [18], where the main conclusion is that if $\check{\check{A}}$ is obtained from the SVD of a perturbed Hankel matrix $\check{\check{H}}(1) = H(1) + E$ and if $\|E\| < \sigma_n(H(1))$ and the dimension of the Hankel matrix is large enough, then the eigenvalue error satisfies

$$|\lambda_j - \tilde{\lambda}_j| \approx \sin \theta, \quad j = 1 : n; \quad (30)$$

where $\sin \theta$ denotes the distance between the subspaces $\mathcal{R}(\mathcal{U}_1)$ and $\mathcal{R}(\tilde{\mathcal{U}}_1)$ with $\tilde{\mathcal{U}}_1$ the matrix of left singular vectors associated with the n largest singular values of $\tilde{H}(1)$. Thus, since A_K and A share the same eigenvalues, it is concluded that the estimate (30) also holds for Kung's method in balanced coordinates, and that excellent eigenvalues estimates are to be expected provided that N is large enough and $\|E\|_1 \ll \sigma_n(H(1))$, as in this case $\sin \theta \approx 0$.

5 Numerical Simulation

The goal here is to illustrate numerically the theory presented in the paper. For this we choose a system having closely spaced eigenvalues and then simulate two well differentiated cases in order to assess the sensitivity of system eigenvalues to inaccuracies.

In the first case we use *exact* data and our goal is to illustrate how nonnormality of the state transition matrix affects the precision of eigenvalues extracted by using finite precision computations. For our illustration we compute Henrici's departure from normality of A_K and A , the eigenvalue condition numbers κ_j , $\tilde{\kappa}_j$, and the eigenvalue error $|\lambda_j - \tilde{\lambda}_j|$, where $\tilde{\lambda}_j$ denotes the computed eigenvalue. All quantities were computed in MATLAB using square Hankel matrices of increasing orders to illustrate the effect of the dimension of the problem over the quantities computed. The data consists of samples of an impulse response (real) signal, h_k , $k = 1 : 1024$, synthesized from an eighty degree of freedom vibrating system, (see [12] for details), whose signal parameters (system eigenvalues λ_j and weights r_j) as well as the eigenvalue separations, are described in Table 1.

Table 1: Signal parameters corresponding to a vibrating system.

j	$z_j, \tilde{z}_j$	$ z_j , \tilde{z}_j $	$r_j, \tilde{r}_j$	δ_j^2
1	$0.9138 \pm 0.4032i$	0.9988	$-0.0013 \pm 0.0004i$	0.00151
2	$0.9908 \pm 0.1262i$	0.9988	$-0.0008 \pm 0.0032i$	0.00010
3	$0.9774 \pm 0.2062i$	0.9989	$-0.0002 \pm 0.0001i$	0.00548
4	$0.9457 \pm 0.3220i$	0.9990	$-0.0001 \pm 0.0001i$	0.00004
5	$0.9923 \pm 0.1162i$	0.9990	$0.0002 \pm 0.0010i$	0.00001
6	$0.9479 \pm 0.3157i$	0.9991	$-0.0003 \pm 0.0009i$	0.00004
7	$0.9596 \pm 0.2780i$	0.9991	$0.0079 \pm 0.0019i$	0.00155
8	$0.8269 \pm 0.5608i$	0.9991	$0.0015 \pm 0.0009i$	0.01991
9	$0.8977 \pm 0.4387i$	0.9991	$-0.0002 \pm 0.0006i$	0.00151

Figures 1-(a) and 1-(b) show that nonnormality and eigenvalue condition numbers strongly depend on the dimension of the problem, and that, as predicted in theory, the state transition matrix in balanced coordinates A_K has in fact better spectral properties than the state transition matrix A . The eigenvalue error is computed by using several implementations of matrix A_K , namely the $\mathcal{U}_1$ -projection version: $A_K = \Sigma_1^{-1/2} \mathcal{U}_1^* C \mathcal{U}_1 \Sigma_1^{1/2}$, the $\mathcal{U}_1$ pseudo-inverse versions: $A_K = \Sigma^{-1/2} \overline{U}^\dagger U \Sigma_1^{1/2}$, and the corresponding counterparts using the right singular matrix $\mathcal{V}_1$. As a result of the computations it was confirmed that except for the projected based versions, all remaining $\mathcal{U}_1$ -based versions using balanced or non balanced coordinates yield essentially the same eigenvalues. A similar comment applies for the eigenvalues obtained from the $\mathcal{V}_1$ -based versions. The conclusion drawn from the results is that the projection-based versions tend to yield estimates of inferior

quality compared with the remaining implementations either using matrix $\mathcal{U}_1$ or matrix $\mathcal{V}_1$. This difference in quality is evident for moderate values of N , as illustrated in Figure 2-(a).

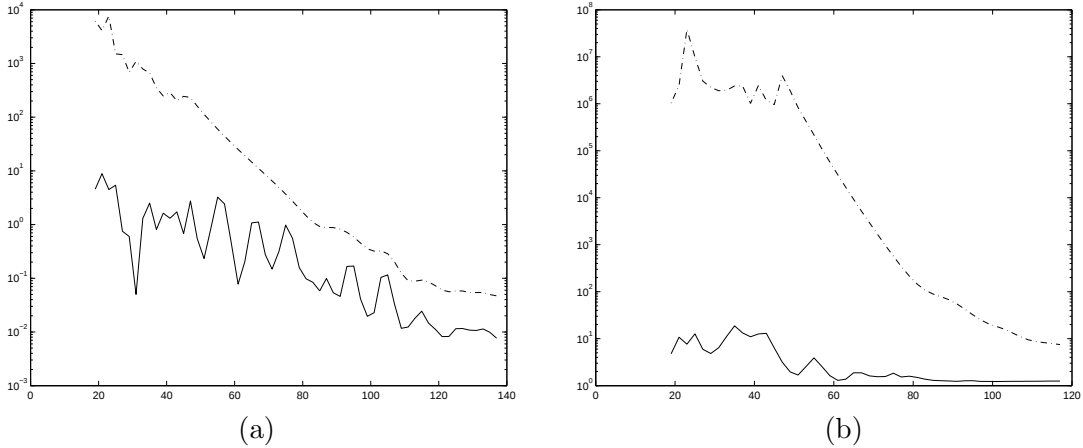


Figure 1: (a): Henrici's departure from normality as a function of N . $D(A_K)^2$ (solid line) and $D(A)^2$ (dash-dotted line). (b): Eigenvalue condition number as a function of N . $\max_j \tilde{\kappa}_j$ (solid line) and $\max_j \kappa_j$ (dash-dotted line). In both cases the Y-axis is in a logarithmic (base 10) scale.

The second case illustrates the performance of Kung's method in recovering system eigenvalues in a noise environment. For this, the exact signal is perturbed by using additive zero mean random noise, adjusted in such a way that $\|\text{noise}\|/\|\text{signal}\| = 0.05$, and then the Kung's matrix eigenvalues are computed. Mean values of 100 noisy realizations are reported. As a result of the computations it is again observed that the quality of the computed eigenvalues by using the projected version is slightly inferior to that obtained from the other versions, see Fig 2-(b). Poor results are because the norm of the error matrix $\|E\|_2 = \|\tilde{H}(1) - H(1)\|_2$ exceeded largely $\sigma_n(H(1))$, i.e., the condition $\|E\|_2 < \sigma_n(H(1))$ that insure the error estimate (30) was no longer valid.

Finally, the spectral stability of balanced and non balanced matrices associated with Kung's method is illustrated. For this, both versions, the balanced and non balanced matrices, were corrupted by random matrices such that $\|\text{state transition matrix}\|/\|\text{random matrix}\|_2 = 0.01$ and then their eigenvalues were extracted. Mean values of eigenvalue error obtained from 100 realizations, displayed in Fig 2-(c), illustrate that the balanced version of Kung's method features slightly better spectral stability than the non balanced one .

6 Conclusion

In this paper spectral properties of state transition matrices of certain projection-type realization methods were investigated based on the estimation of Henrici's departure from normality. It was shown that the system matrix of one of these methods is diagonally similar to the system matrix of Kung's method and proved that the spectral sensitivity of all methods decreases significantly under conditions that are reasonable in practice. Specifically, it was proved that if the Hankel matrix is large enough and the eigenvalues are reasonably close to the unit circle but not extremely close to each other, then the eigenvalues become relatively insensitive to small perturbations, and that in this case the projection-type methods feature essentially the same spectral properties as Kung's method. Further, because both matrices in balanced and non balanced coordinates relate to each

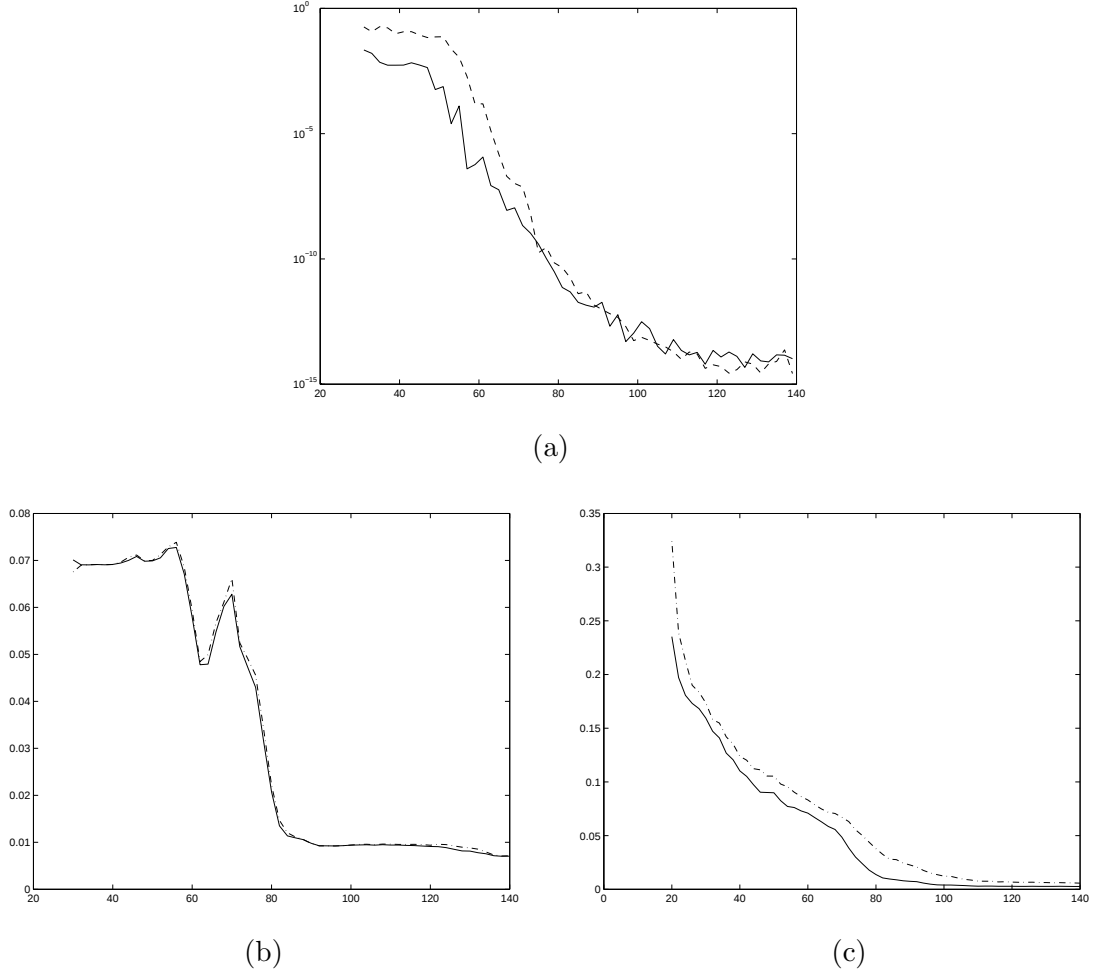


Figure 2: (a) Eigenvalue error as a function of N arising from considering the effect of finite precision computations on “exact data”. Results from projected $\mathcal{V}_1$ - based versions (dashed line). Results from balanced pseudo-inverse $\mathcal{U}_1$ -based versions (solid line). (b) Eigenvalue error as a function of N using additive noise in the signal h_k . Results from balanced $\mathcal{U}_1$ implementation (solid line). Results from non balanced $\mathcal{U}_1$ implementation (dash-dotted line). (c) Eigenvalue error as a function of N using additive noise in the transition matrices. Results from balanced $\mathcal{U}_1$ implementation (solid line). Results from non balanced $\mathcal{U}_1$ implementation (dash-dotted line).

other via a diagonal similarity transformation, it was concluded that eigenvalue error bounds, that apply for state transition matrices in non balanced coordinates, also hold for the balanced case [18]. Experiments were provided to illustrate that the eigenvalue estimates obtained from the balanced case do not outperform so much those obtained from the non balanced ones (Fig. 2-b). The analysis of nonnormality and results of conditioning may be useful in system analysis when modifications of the state transition matrix are needed. Further research for assessing the spectral properties of system matrices in the multi-input multi-output case is needed.

Acknowledgments. The author wishes to thank I. S. Duff and members of the ALGO Team at CERFACS for providing a cordial environment.

References

- [1] Gray SW, Verriest EI. On the sensitivity of generalized state-space systems. *Proceedings of the 28th IEEE Conference on Decision and Control*, Tampa, Florida, 1989, vol.2 pp.1337-1342.
- [2] Kung SY. A new identification and model reduction algorithm via singular value decomposition. *Proceedings of the 12th Asilomar Conference on Circuits, Systems and Computers*, Pacific Grove, 1978, CA, pp.705-714.
- [3] Kung SY, Arun KS, Bhaskar Rao DV. State-space and singular value decomposition-based approximation methods for the harmonic retrieval problem. *Journal of the Optical Society of America* 1983; vol. 73, no. 12, pp. 1799-1811.
- [4] Moore B. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control* 1981; AC-26 No. 1, pp. 17-32.
- [5] Glover K. All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ -error bounds. *International Journal of Control* 1984; 39: 11515-1193.
- [6] Glover K, Curtain RF, Partington JR. Realization and approximation of linear infinite-dimensional systems with error bounds. *SIAM Journal on Control and Optimization* 1988; 16(4): 863-898.
- [7] Lam K, Anderson BDO. Li-impulse response error bound for balanced truncation. *Systems and Control Letters* 1992; 18:129-137.
- [8] Lam J, Hung SY. On the stability of Projections of Balanced Realizations. *Linear Algebra and its Applications* 1987; 257:163-182.
- [9] Li F, Vaccaro RJ. Unified analysis for DOA estimation algorithms in array signal processing. *Signal Processing* 1991; 25:147-169.
- [10] Rao BD. Relationship between matrix pencil and state-space based harmonic retrieval methods. *IEEE Transactions On Acoustics, Speech and Signal Processing* 1990; Vol. 38, No. 1, 177-179.
- [11] Van der Veen A, Deprettere EF, Swindlehurst AL. Subspace-Based Signal Analysis Using Singular Value Decomposition. *Proceedings of the IEEE* 1993; 81(9):1277-1309.
- [12] Bazán FSV, Bavastri C. An Optimized Pseudo-Inverse Algorithm (OPIA) For Multi-Input Multi-Output Modal Parameter Identification. *Mechanical Systems and Signal Processing* 1996; 10:365-380.
- [13] Van Huffel S, Chen H, Decanniere C, Van Hecke P. Algorithm for time-domain NMR data fitting based on Total least squares. *Journal of Magnetic Resonance* 1994; 110:228-237.
- [14] Chaitin-Chatelin F, Frayssé V. Lectures on Finite Precision Computations, SIAM Philadelphia 1996.
- [15] Osborne EE. On pre-conditioning of matrices. *Journal of the Association for Computing Machinery (ACM)* 1960; vol. 7, 338-345.

- [16] Golub GH, Van Loan CF. Matrix Computations. Johns Hopkins University Press, Baltimore, third edition, 1996.
- [17] Bazán FSV. Conditioning of Rectangular Vandermonde Matrices with nodes in the Unit Disk. *SIAM Journal on Matrix Analysis and Applications* 2000; 21(2) pp. 679-693.
- [18] Bazán FSV. Error analysis of signal zeros: a projected companion matrix approach. *Linear Algebra and its Applications* 2003; 369, 153-167.
- [19] Wilkinson JH. The Algebraic Eigenvalue Problem. Oxford University Press, Oxford, Great Britain, 1965.
- [20] Bazán FSV, Toint PhL. Conditioning of Infinite Hankel Matrices of finite rank. *Systems and Control Letters* 2000; 41:347-359.
- [21] Smith RA. The condition numbers of the matrix eigenvalue problem. *Numerische Mathematik* 1967; vol. 10, 232-240.