Resonant Mode Extraction from Scattered Noisy Data*

S. Unnikrishna Pillai and Hyun Seok Oh
Department of Electrical Engineering
Polytechnic University
Five Metrotech Center
Brooklyn, New York 11201
Tel: (718)-260-3732; FAX: (718)-260-3906
e-mail: pillai@arma.poly.edu, hyun@arma.poly.edu

Abstract

This paper presents a technique that combines Prony's original approach together with the high resolution eigenstructure based methods to estimate resonant modes from scattered noisy data. The problem is equivalent to obtaining best rational approximations that fit the actual measurements in a least square sense.

1 Introduction

The problem of identifying a linear time-invariant (LTI) systems from measurements of their output responses to a known input excitation, such as a white noise source or an impulse function, is of fundamental importance in many areas in engineering. Such an identification naturally allows one to predict the system outputs and their resonant modes, and, as a result, this problem has considerable impact in many areas.

Physical systems such as above may not be always linear or time-invariant. Nevertheless, over a reasonable time-interval, the outputs can be assumed to be generated by an LTI system, and for an accurate description over a long period it may be enough to update the system parameters adaptively. Moreover, the LTI system under consideration may not possess a rational transfer function. In that case it becomes necessary to model such systems with reasonable accuracy through a finite-order stable recursive model that is optimal in some fashion. Hence, within this approach, physical systems can be parametrized accurately through a finite-order recursive model, provided the system output can be used to describe the underlying system in a satisfactory manner.

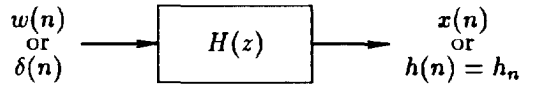

Figure 1: A Linear Time-Invariant (LTI) System

A rational system, as the name implies, possesses a transfer function that is the ratio of two polynomials. In the discrete case, such a system can be compactly described in the z-domain using its transfer function $H(z)$ given by[1]

$$H(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z + \cdots + b_q z^q}{1 + a_1 z + \cdots + a_p z^p}, \qquad (1)$$

where a_i, $i = 1 \rightarrow p$, and b_j, $j = 0 \rightarrow q$ represent the system parameters. Such systems are known as AutoRegressive Moving-Average (ARMA) models with denominator degree equal to p and numerator degree equal to q, or in short, as ARMA(p, q) systems. It is not difficult to grasp their physical meaning. When driven by an input $w(n)$, to generate the output $x(n)$, as shown in Fig. 1, we have

$$x(n) = -\sum_{k=1}^{p} a_k x(n-k) + \sum_{k=0}^{q} b_k w(n-k). \qquad (2)$$

*This research was supported by the Office of Naval Research under contract N00014-89-J-1512 P-5 and the Army Research Office under contract DAAH04-93-02-0010.

[1] We have used the variable z, rather than the usual z^{-1}, for the delay operator. In this representation, a stable function has all its poles outside the unit circle ($|z| > 1$). Minimum phase systems have all their poles and zeros outside the unit circle, and stable nonminimum phase systems have no restrictions on their zeros. The use of the variable z translates all stability arguments to be carried out in the compact region $|z| \leq 1$.

Thus, the present value of the output $x(n)$ depends regressively upon its previous p sample values as well as the running (moving) average generated from $(q+1)$ past samples of the input $w(n)$.

Clearly, if the system is analytic in $|z| < 1$, it can be represented by a one-sided Taylor series expansion given by

$$H(z) = \sum_{k=0}^{\infty} h_k z^k, \quad |z| < 1. \tag{3}$$

Here, $\{h_k\}_{k=0}^{\infty}$ denotes the impulse response sequence that determines the intrinsic characteristics about the system. This sequence can be observed when the system is excited by an impulse function $\delta(n)$. If the input process is a wide-sense stationary white noise process $w(n)$ with power spectral density $S_w(\theta) = \sigma^2$, then the power spectral density of the output process $x(n)$ is given by

$$S_x(\theta) = S_w(\theta)|H(e^{j\theta})|^2 = \sigma^2|H(e^{j\theta})|^2 = \sum_{k=-\infty}^{+\infty} r_k e^{jk\theta}, \tag{4}$$

where

$$r_k \triangleq E[x(n)\, x^*(n+k)], \quad k = 0, 1, \ldots, \infty \tag{5}$$

represents the kth autocorrelation term of the wide-sense stationary output process $x(n)$. If the system is exited by a white noise process $w(n)$, the output correlation r_k can be expressed in terms of the system impulse response h_k as

$$r_k = E[x(n)\, x^*(n+k)] = \sum_{i=0}^{\infty} h_i\, h_{i+k}^* \tag{6}$$

If autocorrelations alone are used in identification problems, then as is well known, systems can be identified only upto their minimum phase equivalent parts [1]. Since a good majority of physical systems are not minimum phase, the above minimum phase factor corresponding to the original system hardly gives the exact phase characteristics of the system. To overcome this difficulty, instead of using only the output autocorrelations r_k, we may need to use additional information such as the first-order or higher-order moments or their combinations. In the next section, we address this problem using partial information regarding the impulse response sequence.

2 Rational System Modeling

A rational system $H(z)$ such as in (1) can be alternatively represented as

$$H(z) = \sum_{k=1}^{p} \frac{c_k}{1 - z_k z}, \quad |z| < 1 \tag{7}$$

where $z_1^{-1}, z_2^{-1}, \cdots z_p^{-1}$ represent the poles of the system.[2] Notice that (see footnote 1) for analyticity of $H(z)$ in $|z| < 1$, the poles must be outside the unit circle, or , $|z_k| < 1$, $k = 1 \rightarrow p$. Thus, using (3), (7) gives

$$h_n = \begin{cases} \sum_{k=1}^{p} c_k\, z_k^n, & n \geq 0, \quad |z_k| < 1 \\ 0, & \text{otherwise} \end{cases}. \tag{8}$$

From (8), given a finite set of the impulse response, the identification problem reduces to evaluating c_k, z_k, $k = 1 \rightarrow p$. Notice that z_k, $k = 1 \rightarrow p$ appear in (8) in a nonlinear manner, and hence direct evaluation of these quantities are in general difficult. Moreover, exact modeling is valid only when the orignal system happens to be representable as a linear combination of p exponential modes, i.e., the orignal system is rational. If the impulse response sequence originated from a nonrational system, to obtain a rational approximation such as above, a measure of 'goodness of fit' must be assumed, so that the unknown parameters can be selected in a systematic fashion by finding the 'best' approximation. Various techniques have been developed to solve this problem such as those based on the Singular Value Decomposition (SVD), the maximum likelihood (ML), the polynomial method and the matrix pencil technique [2, 3].

As is well known, Prony has recognized this problem and ingeniously developed an alternate approach where the nonlinearity in (8) is turned into a linear problem by exploiting the finite order of a rational system [4]. To see this approach, it is best to make use of (1) and (3) to derive a set of equalities that can be used to obtain the desired linear equations. Towards this, equating (1) and (3), and comparing coefficients of like powers, we get

$$b_k = a_0 h_k + a_1 h_{k-1} + \cdots + a_k h_0, \quad k = 0 \rightarrow q \tag{9}$$

and

$$0 = a_0 h_{r+p} + a_1 h_{r+p-1} + a_2 h_{r+p-2} + \cdots + a_p h_r, \quad r \geq 0. \tag{10}$$

Notice that (10) represents an infinite set of linear equations in the unknown quantities a_1, a_2, $\cdots a_p$, and the first p equations there can be used to evaluate

[2] z_k, $k = 1 \rightarrow p$ does not have to be distinct. For simplicity, we will only consider the distinct root situation in this paper.

these unknowns. This gives

$$\begin{bmatrix} h_0 & h_1 & \cdots & h_{p-1} \\ h_1 & h_2 & \cdots & h_p \\ h_2 & h_3 & \cdots & h_{p+1} \\ \vdots & \vdots & \vdots & \vdots \\ h_{p-1} & h_p & \cdots & h_{2p-2} \end{bmatrix} \begin{bmatrix} a_p \\ a_{p-1} \\ a_{p-2} \\ \vdots \\ a_1 \end{bmatrix} = - \begin{bmatrix} h_p \\ h_{p+1} \\ h_{p+2} \\ \vdots \\ h_{2p-1} \end{bmatrix} . \tag{11}$$

In principle, (11) can be used to solve for $a_1, a_2, \cdots a_p$. In that case, from (1) and (7)

$$A(z) = 1+a_1 z+a_2 z^2+\cdots+a_p z^p = (1-z_1 z)\cdots(z-z_p z), \tag{12}$$

and the poles of the system correspond to the zeros of the denominator polynomial $A(z)$. Notice that in Prony's approach, the nonlinear problem of determining z_k, $k = 1 \rightarrow n$ from the impulse response has become a linear problem in (11). Once the a_k's are determined, the numerator coefficients b_k's can be evaluated from (9) directly. Alternatively, since z_k's are known, the residues c_k, $k = 1 \rightarrow p$ in (7), can be determined from the first p equations in (8) as

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ z_1 & z_2 & \cdots & z_{p-1} & z_p \\ z_1^2 & z_2^2 & \cdots & z_{p-1}^2 & z_p^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1^{p-1} & z_2^{p-1} & \cdots & z_{p-1}^{p-1} & z_p^{p-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_p \end{bmatrix} = \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{p-1} \end{bmatrix} . \tag{13}$$

If $z_k \neq z_i$ for $i \neq j$, the Vandermonde matrix in (13) is nonsingular, and a unique solution is guaranteed for the unknown residues $c_1, c_2, \cdots c_p$.

The Prony equations in (10) in fact says much more about the structure of the rational system in (1). To see this, define the Hankel matrices

$$\mathbf{H}_k \triangleq \begin{bmatrix} h_0 & h_1 & \cdots & h_k \\ h_1 & h_2 & \cdots & h_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ h_k & h_{k+1} & \cdots & h_{2k} \end{bmatrix},$$

$$\mathbf{C}_k \triangleq \begin{bmatrix} h_1 & h_2 & \cdots & h_k \\ h_2 & h_3 & \cdots & h_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ h_k & h_{k+1} & \cdots & h_{2k-1} \end{bmatrix}, \quad k \geq 0 . \tag{14}$$

Then from (11), $\mathbf{H}_p$ is singular and $\mathbf{H}_{p-1}$ and $\mathbf{C}_p$ are nonsigular. In fact,

$$rank\, \mathbf{H}_p = rank\, \mathbf{H}_{p-1} = rank \mathbf{C}_p = p . \tag{15}$$

Moreover, from the Prony equation in (10)

$$h_{p+r} = -(a_1 h_{p+r-1} + a_2 h_{p+k-2} + \cdots + a_p h_r) , \quad r \geq 0 \tag{16}$$

and consequently if $k \geq p$, rows/columns beyond the p^{th} row/column in $\mathbf{H}_k$ and $\mathbf{C}_k$ are linearly dependent on their previous rows/columns. Thus, together with (15), it follows that

$$rank\, \mathbf{H}_{k-1} = rank\, \mathbf{C}_k = p , \quad k \geq p . \tag{17}$$

Equation (17) shows the rich structure present in a rational system, and it forms the necessary and sufficient conditions for an infinite sequence $\{h_k\}_{k=0}^{\infty}$ to represent the impulse response sequence of a rational system. In principle, the rank condition (17) can be used to determine the model order p of the rational system, and then Prony's approach can be successfully applied to evaluate the system parameters as described before.

3 Noisy Data Case

The situation is however not so ideal if the impulse response data h_n, $n \geq 0$ is corrupted by noise. In that case, the rank conditions in (17) are no longer true, and it will be difficult to determine the model order using that approach. Since measured data is always noisy, we may model the observed data as

$$x(n) = h_n + w(n) = \sum_{k=1}^{p} c_k z_k^n + w(n) , \tag{18}$$

where $w(n)$ represents additive noise that corrupts the impulse response. In this case, the problem is statistical, and the principle of maximum likelihood (ML) may be used to evaluate the unknown nonrandom parameters c_k, z_k, $k = 1 \rightarrow p$ in (18). The ML approach first computes the joint probability density function of the observations, and then evaluates the log-likelihood function. In principle, the optimal values of the unknowns correspond to the peaks of this function, and they can be estimated by a search procedure. However for any reasonable value of p, this nonlinear search is quite tedious, and alternate suboptimal methods must be developed to obtain good solutions to this problem.

Towards this, once again we can make use of the Hankel matrix $\mathbf{H}_k$ defined in (14). Suppose $h_0 \rightarrow h_{2n}$ are available, and let $\mathbf{H} = \mathbf{H}_n$ represent the symmetric Hankel matrix of size $(n+1)\times(n+1)$ as in (14). Since, $\mathbf{H}$ is real and symmetric, we also have $\mathbf{H} = \mathbf{H}^*$, where $\mathbf{H}^* = \overline{\mathbf{H}}^T$ represents the complex conjugate transpose of $\mathbf{H}$. To incorporate the structure of h_k in (8), define

$$\mathbf{Z} \triangleq \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ z_1 & z_2 & \cdots & z_{p-1} & z_p \\ z_1^2 & z_2^2 & \cdots & z_{p-1}^2 & z_p^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1^n & z_2^n & \cdots & z_{p-1}^n & z_p^n, \end{bmatrix} \tag{19}$$

and

$$\mathbf{C} = diag\,[c_1, c_2, \cdots, c_p], \tag{20}$$

where $diag\,[\cdot]$ represents a diagonal matrix. Then using (19)–(20), a direct computation shows that

$$\mathbf{H} = \mathbf{Z}\mathbf{C}\mathbf{Z}^T \tag{21}$$

and with the help of $\mathbf{H} = \mathbf{H}^*$, we also have

$$\mathbf{H} = \overline{\mathbf{Z}}\mathbf{C}^*\mathbf{Z}^* . \tag{22}$$

Notice that $\mathbf{H}$ is of size $(n+1) \times (n+1)$ and rank p. Hence, $n+1-p$ eigenvalues of $\mathbf{H}$ are zeros. Since $\mathbf{H}$ is symmetric, the corresponding eigenvectors are linearly independent and can be chosen to be orthogonal. Let $\mathbf{e}_1, \mathbf{e}_2, \cdots \mathbf{e}_{n+1}$ represent the eigenvectors and $\lambda_1, \lambda_2, \cdots \lambda_{n+1}$ represent the eigenvalues of $\mathbf{H}$. By rearrangement, we can always write

$$\lambda_k \equiv 0\,, \quad k \geq p+1\,. \tag{23}$$

Thus

$$\mathbf{H}\mathbf{e}_k = \lambda_k \mathbf{e}_k\,, \quad k = 1 \rightarrow p\,, \quad |\lambda_k| \neq 0 \tag{24}$$

and

$$\mathbf{H}\mathbf{e}_k = 0\,, \quad k = p+1 \rightarrow n+1\,. \tag{25}$$

Using (22) and (25), this gives

$$\mathbf{Z}^*\mathbf{e}_k = 0\,, \quad k = p+1 \rightarrow n+1\,. \tag{26}$$

Let

$$\mathbf{v}_i = [\,1,\, z_i,\, z_i^2,\, \cdots z_i^n\,]^T\,. \tag{27}$$

Then from (19)

$$\mathbf{Z} = [\,\mathbf{v}_1,\, \mathbf{v}_2,\, \cdots \mathbf{v}_p\,] \tag{28}$$

and using (28) in (26), we get

$$\mathbf{v}_i^*\mathbf{e}_k = 0\,, \quad i = 1 \rightarrow p\,, \quad k = p+1 \rightarrow n+1\,, \tag{29}$$

i.e., the p linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$, in (27) are orthogonal to the eigenvectors $\{\mathbf{e}_{p+1}, \mathbf{e}_{p+2}, \cdots, \mathbf{e}_{n+1}\}$ associated with the zero eigenvalue of $\mathbf{H}$. Thus

$$\{\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_p\} \perp \{\mathbf{e}_{p+1}, \mathbf{e}_{p+2}, \cdots \mathbf{e}_{n+1}\}\,. \tag{30}$$

From (24)–(25), we also have

$$\{\mathbf{e}_1, \mathbf{e}_2, \cdots \mathbf{e}_p\} \perp \{\mathbf{e}_{p+1}, \mathbf{e}_{p+2}, \cdots \mathbf{e}_{n+1}\}\,, \tag{31}$$

since the eigenvectors associated with distinct eigenvalues of a symmetric matrix are orthogonal to each other. Finally using (30) and (31), we get that the set of linearly independent vectors

$$\{\mathbf{e}_1, \mathbf{e}_2, \cdots \mathbf{e}_p\} \quad \text{and} \quad \{\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_p\}$$

span the same subspace of dimension p. As a results,

$$\mathbf{e}_i = \sum_{k=1}^{p} d_{ik}\mathbf{v}_k\,, \tag{32}$$

or

$$\mathbf{E} \triangleq [\mathbf{e}_1, \mathbf{e}_2, \cdots \mathbf{e}_p] = [\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_p]\,\mathbf{D} \tag{33}$$

where $\mathbf{D}_{ik} \triangleq d_{ik}$. Thus,

$$\mathbf{E} = \mathbf{Z}\mathbf{D}\,. \tag{34}$$

Equation (34) can be used to determine the desired unknowns $z_1, z_2, \cdots z_n$. For example, let $\mathbf{E}_1$ and $\mathbf{E}_2$ represent the first and last n rows of $\mathbf{E}$ respectively. Then

$$\mathbf{E}_1 \triangleq \mathbf{Z}_1\mathbf{D} \tag{35}$$

$$\mathbf{E}_2 \triangleq \mathbf{Z}_2\mathbf{D} = \mathbf{Z}_1\mathbf{B}\mathbf{D} \tag{36}$$

where $\mathbf{Z}_1$ abd $\mathbf{Z}_2$ represent two $n \times n$ matrices generated from the first and last n rows of $\mathbf{Z}$ respectively, with $\mathbf{Z}$ is as given in (19). Further

$$\mathbf{B} \triangleq diag\,[z_1, z_2, \cdots, z_p]. \tag{37}$$

From (36)–(37), the generalized eigenvalues of $\mathbf{E}_1$ with respect to $\mathbf{E}_2$ are given by

$$\mathbf{E}_1 - \mu\mathbf{E}_2 = \mathbf{Z}_1\mathbf{D} - \mu\mathbf{Z}_1\mathbf{B}\mathbf{D} = \mathbf{Z}_1\,(\mathbf{I} - \mu\mathbf{B})\,\mathbf{D}\,, \tag{38}$$

or the desired eignevalues are given by

$$\mu_i = z_i^{-1}\,, \quad i = 1 \rightarrow p\,. \tag{39}$$

From (38), since the eigenvalue of $\mathbf{E}_1 - \lambda\mathbf{E}_2$ and $(\mathbf{E}_2^*\mathbf{E}_2)^{-1}\,\mathbf{E}_2^*\mathbf{E}_1 - \mu\mathbf{I}$ are the same, for example, the later form can be used to determine the desired pole locations z_i^{-1}, $i = 1 \rightarrow n$, in (39).

Notice that the above procedure contains a double eigendecomposition procedure, and has been observed to work well in presence of noise. Of course, in the absence of noise, this prcedure is unnecessary, since the simpler Prony's original approach will guarantee the true solutions. However, unlike Prony's approach that determines the denominator coefficients a_k, $k = 1 \rightarrow p$ in (11), the present approach directly determines the pole locations z_k^{-1}, $k = 1 \rightarrow p$, from the data through a double eigendecomposition method. Since,

$$z_k^{-1} = e^{(\sigma_k + j\omega_k)}\,, \tag{40}$$

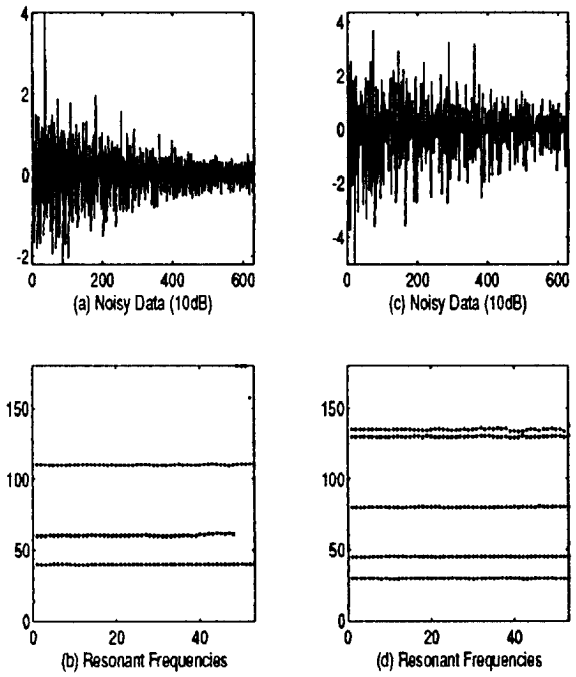

Figure 2: Resonant Frequency extraction of LTI systems from noisy output data

where $\sigma_k > 0$ represents the damping factor and ω_k the resonant frequency, using (40) in (18) it follows that the unknown parameters σ_k and ω_k appear in the observed data in the form of an FM signal in noise, and hence their direct determination from the observed data must be superior to Prony's coefficient determination method which only performs like an AM signal in noise. The noisy data is used in $\mathbf{H} = \mathbf{H}_n$ directly, and the above double eigendecomposition procedure is performed to obtain estimates for the pole locations $\hat{z}_1^{-1}, \hat{z}_2^{-1}, \cdots \hat{z}_p^{-1}$. Notice that determination of the model order p has to be done using statistical procedures on the set of eigenvalues of $\mathbf{H}$, since no clear cut separation as in (24)-(26) will be available in the case of noisy data. The lowest eigenvalues of $\hat{\mathbf{H}}$ may be grouped together to represent small perturbations of the zero eigenvalue of $\mathbf{H}$, and this procedure can be used to obtain a good estimate for the model order p. Since $\hat{\mathbf{H}}$ uses all available data, the two eigendecompositions will smooth out the effect of noise. Once $\hat{z}_k$ are obtained in this manner, the residues $\hat{c}_k$ can be estimated using the least square solution on the whole data set as in (13) by making use of $(2n+1)$ rows there.

Figs 2.(a)-(b) show results of simulation for a degree six system with poles located at $e^{(\sigma_k \pm j\omega_k)}$, $k = 1 \rightarrow 3$ with $\sigma_1 = 0.002$, $\omega_1 = 40°$, $\sigma_2 = 0.006$, $\omega_2 = 60°$, $\sigma_3 = 0.005$ and $\omega_3 = 110°$. Similarly Figs 2.(c)-(d) represent a tenth degree system with poles at $e^{(\sigma_k \pm j\omega_k)}$, $k = 1 \rightarrow 5$ with $\sigma_1 = 0.0015$, $\omega_1 = 30°$, $\sigma_2 = 0.001$, $\omega_2 = 45°$, $\sigma_3 = 0.003$, $\omega_3 = 80°$, $\sigma_4 = 0.002$, $\omega_4 = 130°$, $\sigma_5 = 0.003$ and $\omega_5 = 135°$. In both cases, $10dB$ SNR is maintained by adding noise as in (18) (see Figs 2(a), 2(c)). A sliding window of width equal to 100 data samples generates the noisy data that is processed using the method described in Section III, and the estimated resonant frequencies ω_k, $k = 1, 2, \cdots$ are plotted in Figs 2(b), 2(d), against the window location. The sliding window is sequentially moved to the right in increments of ten data samples. From Figs 2(b), 2(d) it is clear that the new method performs quite well in estimating the resonant frequencies of LTI systems from noisy impulse response data.

4 Conclusions

This paper presents a new approach to resonant mode extraction by combining Prony's original approach together with the high resolution eigenstructure based methods. The rationality of the original system exhibits certain rank invariant Hankel structures, and in the noisy case this is exploited in a least square sense to extract the best possible resonant frequencies of the underlying system.

REFERENCES

[1] S. U. Pillai and T. I. Shim, and D. C. Youla, "A New technique for ARMA-system identification and rational approximation," *IEEE Trans. Signal Processing*, vol.SP-41, no.3, pp. 1281-1304, March 1993.

[2] Y. Hua and T. K. Sarkar, "Matrix Pencil Method for Estimating Parameters of Exponentially Damped/Undamped Sinusoids in Noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.38, no.5, pp.814-824, May 1990.

[3] Y. Hua and T. K. Sarkar, "On SVD for Estimating Generalized Eigenvaluses of Singular Matrix Pencil in Noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.39, no.4, pp.892-900, April 1991.

[4] D. C. Youla, "The Identification of Linear Dynamical Systems from Time-Domain Measurements: A critical Study of Certain Aspects of Prony's Method," Department of Electrical Engineering, Polythechnic Institute of New York, Farmingdale, New York, 11735, September 1978.