BLIND IDENTIFICATION OF SINGLE-INPUT MULTIPLE-OUTPUT POLE-ZERO SYSTEMS *

Gopal T. Venkatesan[1] *Mostafa Kaveh*[1] *Ahmed H. Tewfik*[1] *KevinM.Buckley*[2]

[1] Department of Electrical and Computer Engineering
University of Minnesota, Minneapolis, MN 55455.
[2] Department of Electrical and Computer Engineering
Villanova University, Villanova, PA 19085.

ABSTRACT

In this paper we present a technique for the blind identification of single-input multiple-output (SIMO) pole-zero (PZ) systems using only the second order statistics of the system output data. The system input is treated as an unknown deterministic sequence, and hence, restrictive i.i.d. assumptions on the input sequence are not required. We estimate the poles and zeros of the channels in two steps : 1) estimate product of all permutations of a numerator and a denominator polynomial from two different channels, and 2) extract individual numerator and denominator polynomials for each channel from above estimate. Our technique performs well even with short records of data.

1. INTRODUCTION

The multichannel blind identification (BI) problem arises in a wide variety of engineering applications. The feasibility of identifying nonminimum phase finite impulse response (FIR) channels using only the second order statistics (SOS) of the system multichannel output data was first demonstrated in [1]. Since then, considerable work has been done in this area both in algorithm development and fundamental analysis (see [2] and references therein). However, the infinite impulse response (IIR) case with multichannel PZ systems has been relatively unaddressed. In [4] two techniques, the linear prediction (LP) approach and the generalized subspace (GS) approach, for second order blind *source extraction* for multiple-input multiple-output (MIMO) PZ systems have been presented. The above two approaches address the general problem of MIMO PZ system identification when the number of inputs is strictly less than the number of outputs. They require the sources to be i.i.d., and in cases where source separation is necessary, the sources are assumed to be non-Gaussian.

The LP approach models the output of the irreducible multichannel PZ system as a multichannel autoregressive (AR) system of finite order. Once the parameters of the AR system have been solved for, an estimate of the source sequence is obtained. Note that the parameters of the channels are not estimated using this method. Also, this method requires large blocks of data to provide reasonable correlation estimates. However, the method is quite robust to over-parameterization. The GS approach is based on the canonical right matrix fraction description (MFD) of rational functions [3]. This method first estimates the minimal polynomial basis (MPB) and then extracts the right factor matrix using the LP approach. This subspace approach performs better than the LP approach especially when only short records of data are available. However, it is not robust to model order mismatches.

Our work on the SIMO PZ BI problem was motivated by our investigation into the feasibility of providing reliable fault diagnostics and monitoring using Acoustic Emissions (AE) [6][7]. AEs are ultrasonic waves created due to the formation/propagation of a crack in a material. These waves propagate through the material and can be recorded by piezo-electric sensors placed strategically on the surface. With this recorded data at various sensors the need to identify the propagation characteristics (channel sequence) and crack (input) sequence for subsequent crack localization/characterization arises naturally. This is obviously a SIMO BI problem. However, the situation in this context is quite different from the communication scenario that has motivated research in SIMO BI in the past several years. For communication scenarios BI methods make standard assumptions of FIR channels and/or white, persistent sources. In the fault monitoring case, from experimental data, we have found that the above assumptions are no longer valid. This initiated our investigation into BI of SIMO PZ systems with a non-white input sequence. In this paper we deal only with the persistent input sequence case for SIMO PZ systems. A method for SIMO PZ BI when the input sequence is of finite length is discussed in [8].

Our interest focuses on SIMO systems. Though the SIMO problem can be addressed by the above methods, the structure of the single input case enables us to use a simpler approach that can provide both the channel and input sequence estimates. The least squares (LS) method presented in this paper takes advantage of the commutativity property of the convolution operation. A LS solution in the SIMO FIR case using the same property has been presented in [5]. In our method the input is treated as an unknown deterministic sequence, and in consequence, unlike existing methods for multichannel PZ system identification it is not limited to i.i.d. input scenarios. In this paper, we assume prior knowledge of the model orders. Our technique is sensitive to model order mismatches especially in the context of the channel parameter estimates. However, reasonable robustness to model order mismatches in terms of input sequence estimates has been observed in our simulations.

2. THE PROPOSED LS TECHNIQUE

2.1. The SIMO PZ Model

Consider an M-channel SIMO system whose channel transfer functions are rational, i.e., the z-transform of the impulse response of the i^{th} channel is given by,

$$\begin{aligned} H_i(z) &= \frac{B_i(z)}{A_i(z)} \\ &= \frac{b_i(0) + b_i(1)z^{-1} + \ldots + b_i(q_i - 1)z^{-q_i+1}}{1 + a_i(1)z^{-1} + \ldots + a_i(p_i - 1)z^{-p_i+1}} \end{aligned} \quad (1)$$

*THIS WORK WAS SUPPORTED BY THE OFFICE OF NAVAL RESEARCH UNDER MURI CONTRACT # N00014-95-1-0539.

where $i = 0, 1, 2...M-1$. The system output in the z-domain is given by,

$$\mathbf{Y}(z) = \mathbf{H}(z)S(z) \tag{2}$$

where

$$\mathbf{Y}(z) = \begin{pmatrix} Y_0(z) \\ Y_1(z) \\ \vdots \\ Y_{M-1}(z) \end{pmatrix} \qquad \mathbf{H}(z) = \begin{pmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{M-1}(z) \end{pmatrix}. \tag{3}$$

Also, S(z) is the common input to the different channels and is a scalar polynomial. Now consider two of the channels in (2). Using (1) we get,

$$A_i(z)Y_i(z) = B_i(z)S(z) \qquad A_j(z)Y_j(z) = B_j(z)S(z). \tag{4}$$

Multiplying both sides of the i^{th} channel equation by $B_j(z)$ and both sides of the j^{th} channel equation by $B_i(z)$, and subtracting we get,

$$B_j(z)A_i(z)Y_i(z) - B_i(z)A_j(z)Y_j(z) = 0. \tag{5}$$

Note that the commutativity property of the convolution operation has been used. Let

$$C_{ij}(z) = B_j(z)A_i(z) \tag{6}$$

where $i, j = 0, 1, ..., M-1$ and $i \neq j$. From (5) we can see that we can solve for all $K = M(M-1)$ combinations of $C_{ij}(z)$ in a straightforward way. However, the recovery of $B_j(z)$ and $A_i(z)$ from $C_{ij}(z)$ is not possible, for example when $M = 2$, as there is no possible way of distinguishing between the poles and zeros in (5). But it can be shown that the poles and zeros can be separated without any ambiguity when there are 3 or more channels. We will discuss this result further after we have couched the above equations in matrix-vector notation.

2.2. Identifying the product $B_j(z)A_i(z)$

Let P and Q be the maximum values of the $p_i's$ and $q_i's$ respectively. Hence, $R = P+Q-1$ will be the maximum of the orders of $C_{ij}(z)$. Consider N points of data recorded at each system output. The solutions for a given pair $C_{ij}(z)$ and $C_{ji}(z)$ are coupled as in (5) and the convolution version in the time domain can be represented as a set of linear equations as,

$$\begin{pmatrix} -\mathbf{Y_i} & \mathbf{Y_j} \end{pmatrix} \begin{pmatrix} \mathbf{c_{ij}} \\ \mathbf{c_{ji}} \end{pmatrix} = \mathbf{0_{(N-R+1)\times 1}} \tag{7}$$

where $\mathbf{c_{ij}}$ and $\mathbf{c_{ji}}$ are $R \times 1$ unknown vectors whose elements are the coefficients of the polynomials $C_{ij}(z)$ and $C_{ji}(z)$ respectively. The $(N-R+1) \times 1$ matrix $\mathbf{Y_i}$ is given by,

$$\mathbf{Y_i} = \begin{pmatrix} y_i(R-1) & y_i(R-2) & \cdots & y_i(0) \\ y_i(R) & y_i(R-1) & \cdots & y_i(1) \\ \vdots & \vdots & \vdots & \vdots \\ y_i(N-1) & y_i(N-2) & \cdots & y_i(N-R) \end{pmatrix}. \tag{8}$$

Note that the unknown vector in (7) should be constrained to avoid trivial solutions. Also, for (7) to have a unique solution, the number of equations has to be at least equal to the number of unknowns. Taking into account the contamination of the output observations by additive noise this system of overdetermined equations leads naturally to a LS solution. The above condition can be expressed as $N \geq (3R-1)$ without taking into account the reduced dimensionality of the problem due to the constraint set. For a given system of M channels all combinations of (5) can be formulated as $K/2$ sets of equations. These sets of equations provide unique solutions for all $C_{ij}(z)$ $(i, j = 0, 1, ..., M-1; i \neq j)$.

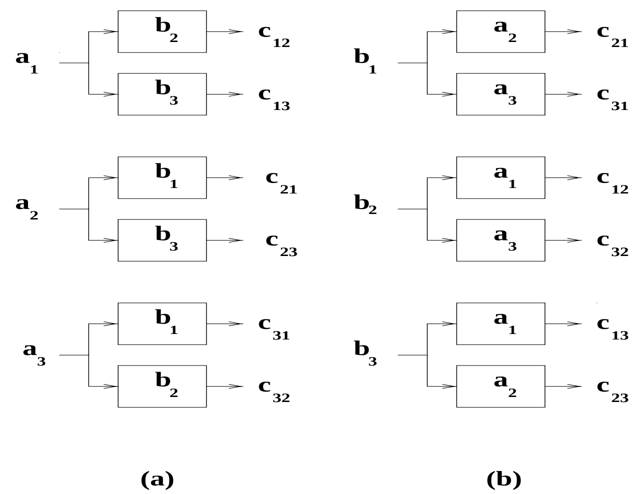

Figure 1. Special MIMO FIR model

2.3. Solving for $A_i(z)$ and $B_j(z)$

Let $\mathbf{a_i}$ and $\mathbf{b_j}$ be the coefficient vectors of the polynomials $A_i(z)$ and $B_j(z)$ respectively. Since $\mathbf{c_{ij}} = \mathbf{a_i} \otimes \mathbf{b_j}$ ($\otimes$ represents the convolution operation) we have a special type of the multichannel FIR BI problem. Figure 1 depicts the 3-channel case [1]. The $MP \times 1$ and $MQ \times 1$ coefficient vectors $\mathbf{a}$ and $\mathbf{b}$ of the multichannel system are given by concatenating the respective coefficient vectors together. Model (a) treats the unknown coefficient vector $\mathbf{a}$ as an unknown "input sequence" propagating through various "channels" given by $\mathbf{b}$. Model (b) reverses their roles. Due to the commutativity of the convolution operation it may seem that using either of the two models shown in Figure 1 would be acceptable. However, this is not the case. The choice is dictated by the values of M, P and Q. Depending on their values only one of the above models may provide a unique solution at this first step. The BI problem depicted in Figure 1 is solved in two steps. Proceeding as per traditional solutions to the multichannel FIR problem [5] the "input sequence" is eliminated and a linear system of equations in the unknown "channel sequence" is first obtained. The answer to whether this system of equations is *not* underdetermined, is intimately related to the values of M, P and Q. In fact, as will be shown later, picking the lower length coefficient vector, $\mathbf{a}$ or $\mathbf{b}$, as the "channel" sequence is a sufficient (not always necessary!) requirement to obtain an overdetermined system of equations.

First consider model (a) for the 3-channel case ($K = 6$). It can be represented as,

$$\mathbf{Ba} = \mathbf{c} \tag{9}$$

where the $6R \times 1$ vector $\mathbf{c} = \begin{pmatrix} \mathbf{c_{31}^T} & \mathbf{c_{13}^T} & \mathbf{c_{32}^T} & \mathbf{c_{23}^T} & \mathbf{c_{21}^T} & \mathbf{c_{12}^T} \end{pmatrix}^{\mathbf{T}}$. The matrix $KR \times MP$ matrix B for the 3-channel case is given by,

$$\mathbf{B} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{B_1} \\ \mathbf{B_3} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B_2} \\ \mathbf{0} & \mathbf{B_3} & \mathbf{0} \\ \mathbf{0} & \mathbf{B_1} & \mathbf{0} \\ \mathbf{B_2} & \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{10}$$

where

$$\mathbf{B_i} = \mathcal{T}_{\mathbf{R\times P}}(\mathbf{b_i}, .). \tag{11}$$

[1] We consider the 3-channel case for simplicity. It is straightforward to extend the structure for larger number of channels

The operation $\mathcal{T}_{RW\times CL}(\mathbf{colv},\mathbf{rowv})$ produces a Toeplitz matrix of size $RW \times CL$ whose first column and row are given by the vectors **colv** and **rowv** respectively. If the required dimensions of the matrix are bigger than the dimensions of the vector parameters then an appropriate number of zeros are appended to the vectors. If the required dimensions of the matrix are smaller than the dimensions of the vector parameters, then the first RW rows and first CL columns of the matrix are taken. Note that the first elements of **colv** and **rowv** should be the same. If either **colv** or **rowv** is zero except for its first element that should be equal to the first element of the other vector which is nonzero we indicate it simply by a period.

Proceeding as in [5] we get,

$$\mathbf{Cb} = \mathbf{0} \tag{12}$$

where for the $M = 3$ case $\mathbf{C}$ is given by,

$$\mathbf{C} = \begin{pmatrix} \mathbf{0} & -\mathbf{C_{13}} & \mathbf{C_{12}} \\ -\mathbf{C_{23}} & \mathbf{0} & \mathbf{C_{21}} \\ \mathbf{C_{32}} & -\mathbf{C_{31}} & 0 \end{pmatrix}. \tag{13}$$

In the above equation $\mathbf{C_{ij}}$ is given by $\mathcal{T}_{P\times Q}(\mathbf{c_{ij}},.)$ as defined in (11).

For the $M = 4$ case $\mathbf{C}$ is given by,

$$\mathbf{C} = \begin{pmatrix} \mathbf{0} & -\mathbf{C_{13}} & \mathbf{C_{12}} & \mathbf{0} \\ \mathbf{0} & -\mathbf{C_{14}} & \mathbf{0} & \mathbf{C_{12}} \\ \mathbf{0} & \mathbf{0} & -\mathbf{C_{14}} & \mathbf{C_{13}} \\ -\mathbf{C_{23}} & \mathbf{0} & \mathbf{C_{21}} & \mathbf{0} \\ -\mathbf{C_{24}} & \mathbf{0} & \mathbf{0} & \mathbf{C_{21}} \\ \mathbf{0} & \mathbf{0} & \mathbf{C_{24}} & \mathbf{C_{23}} \\ -\mathbf{C_{32}} & \mathbf{C_{31}} & \mathbf{0} & \mathbf{0} \\ -\mathbf{C_{34}} & \mathbf{0} & \mathbf{0} & \mathbf{C_{31}} \\ \mathbf{0} & \mathbf{C_{34}} & \mathbf{0} & \mathbf{C_{32}} \\ -\mathbf{C_{42}} & \mathbf{C_{41}} & \mathbf{0} & \mathbf{0} \\ -\mathbf{C_{43}} & \mathbf{0} & \mathbf{C_{41}} & \mathbf{0} \\ \mathbf{0} & -\mathbf{C_{43}} & \mathbf{C_{42}} & \mathbf{0} \end{pmatrix}. \tag{14}$$

Note that for a given M, the size of the $\mathbf{C}$ matrix is $M(M-1)(M-2)P/2 \times MQ$. Now for an overdetermined system of equations we require $M(M-1)(M-2)P/2 \geq MQ$. This in turn implies that $Q \leq (M-1)(M-2)P/2$. Once $\mathbf{b}$ is obtained from (12) (for this model, co-primeness of the numerator polynomials is a sufficient condition to ensure uniqueness of the solution up to a scalar multiple) subject to constraints that preclude a trivial solution, $\mathbf{a}$ can be obtained using (9).

Alternatively, if we had used the model shown in Figure 1(b) the size of the $\mathbf{C}$ matrix for a given M would be $M(M-1)(M-2)Q/2 \times MP$ and the corresponding requirement to obtain an overdetermined system of equations would have been $P \leq (M-1)(M-2)Q/2$ (for this model, co-primeness of the denominator polynomials is a sufficient condition to ensure uniqueness of the solution up to a scalar multiple). These inequalities give a criterion for choosing between the two models shown in Figure 1 depending on the number of channels. For example, when $M = 3$ we have, $Q \leq P$ for model (a) and $P \leq Q$ for model(b). Hence, the choice of the appropriate model will depend upon whether the denominator order P is greater than the numerator order Q or vice versa. Now if $M = 4$ then the condition for using model (a) is given by $Q \leq 3P$ and so on.

Also, note that from the structure of $\mathbf{C}$ *for* $M = 2$ *we can never obtain an overdetermined system of equations for any value of* P *and* Q. *This makes it clear why the problem cannot be solved for the 2-channel case without ambiguity.*

Once the channel parameters have been found it is straightforward to obtain the unknown source sequence. Rewriting the system in matrix-vector notation as,

$$\bar{\mathbf{A}}\mathbf{y} = \bar{\mathbf{B}}\mathbf{s} \tag{15}$$

where $\mathbf{y} = [\mathbf{y_0^T} \ldots \mathbf{y_{M-1}^T}]^\mathbf{T}$, the M, $N \times 1$ vectors $\mathbf{y_i}$ being the N points of data observed at the output of the corresponding channel and $\mathbf{s}$ is the unknown $(N+Q-1) \times 1$ source vector. The $MN \times (N+Q-1)$ matrix $\bar{\mathbf{B}}$ represents the moving average (MA) operation of the transfer function and is given by,

$$\bar{\mathbf{B}} = \begin{pmatrix} \bar{\mathbf{B}}_\mathbf{0} \\ \bar{\mathbf{B}}_\mathbf{1} \\ \vdots \\ \bar{\mathbf{B}}_\mathbf{M-1} \end{pmatrix} \tag{16}$$

where $\bar{\mathbf{B}}_\mathbf{i} = \mathcal{T}_{\mathbf{N\times N}}(\mathbf{b_i},.)$ as defined in (11). Note that in the fault monitoring application we are interested in, transient signals are being monitored. Hence, the output can be assumed to be zero for time instants outside the duration of the transient. Therefore, if the observation window is the duration of the transient, (15) is an exact representation of the system without "end-effects" even though only a finite amount of data is available. The $MN \times MN$ matrix $\bar{\mathbf{A}}$ in (15) is a block diagonal matrix with $N \times N$ matrices $\mathbf{A_i} = \mathcal{T}_{\mathbf{N\times N}}(\mathbf{a_i},.)$ along the diagonal. From (15) the source vector estimate is given by,

$$\mathbf{s} = \bar{\mathbf{B}}^\dagger \bar{\mathbf{A}}\mathbf{y} \tag{17}$$

where '†' denotes the pseudo-inverse of the matrix.

Summarizing the procedure we have,

Step I : Check if the number of collected data points per channel N is more than the minimum required to provide a unique solution for $\mathbf{c}$.

Step II : Solve for all $\mathbf{c_{ij}}$ using (7).

Step III : Choose an appropriate model (either (a) or (b) shown in Figure 1) keeping in mind the conditions on P and Q for a given value of M.

Step IV : Solve for $\mathbf{a}$ and $\mathbf{b}$ in the appropriate order.

Step V : Solve for the input sequence $\mathbf{s}$ using (17)

Note that no assumptions have been made on the characteristics of the input sequence. The input vector is assumed to be deterministic and can be a segment/realization of a deterministic/stochastic signal.

3. SIMULATIONS

In all the following simulations we assume that $a_i(0) = 1$ and that $b_i(0)$ is known for all the channels. This provides a simple linear constraint $c_{ij}(0) = a_i(0)b_j(0)$ when solving for $\mathbf{c_{ij}}$ in (7).

As mentioned earlier one of the advantages of our approach is its applicability to non-white input cases. We will first consider such a scenario with a 3-channel system with parameters $P = 3$, $Q = 7$ and $N = 60$. Since $P < Q$, model (b) is an obvious choice. Figure 2 shows the actual input sequence and its magnitude spectrum (solid line) superimposed on the estimates obtained with our method (dashed line) at an SNR of 30 dB. Note that the methods outlined in [4] are not applicable in this case due to non-whiteness of the input. Figure 3 shows the source estimate obtained at an SNR of 60 dB for the above situation using the LP method.

Our method can also be applied to communication situations without any modifications. Consider again a 3-channel system with parameters $P = 3$, $Q = 7$ and $N = 60$. The signal constellations after equalization at an SNR of 35dB using our method and the LP method are shown in Figures 4 and 5 respectively (a superimposition of 15 realizations are shown). Note that our approach performs well even with a short record of data. There was no symbol error after equalization.

Above we have presented only input sequence estimates. Channel parameter estimates are also provided by our technique as an intermediate step.

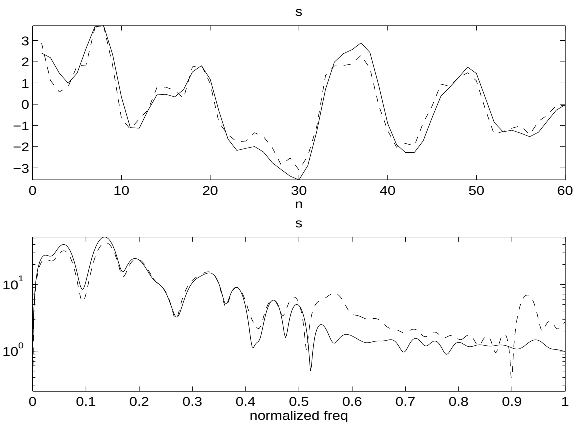

Figure 2. Source Estimates for SNR=30 dB

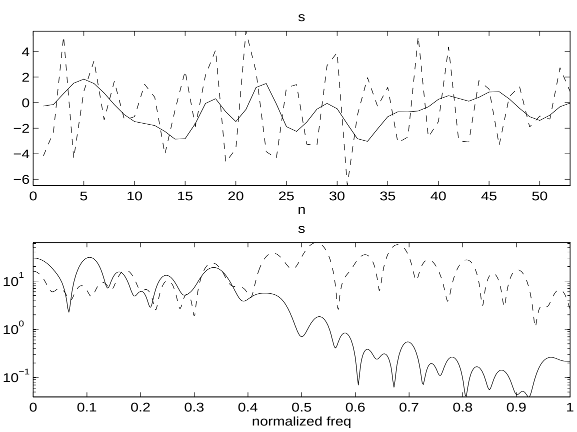

Figure 3. Source Estimates for SNR=60 dB (LP method)

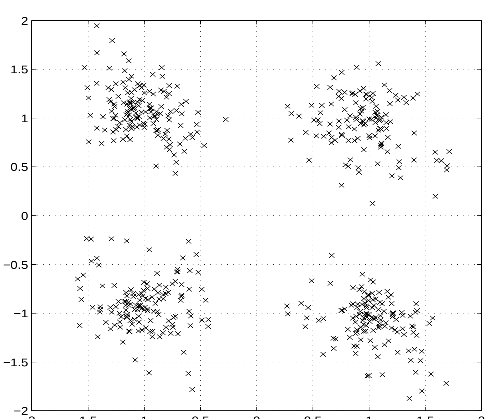

Figure 4. Signal Constellation after Equalization (The paper's method), $N = 60$

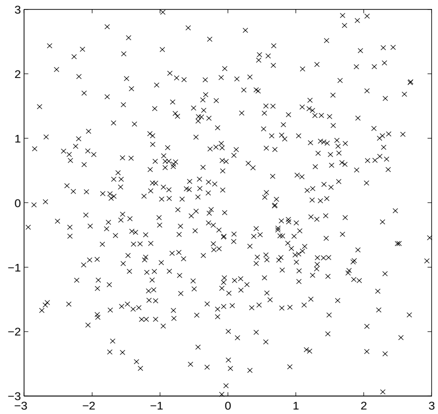

Figure 5. Signal Constellation after Equalization (LP method), $N = 60$

4. CONCLUSIONS

We have presented a novel approach to the blind identification of SIMO PZ systems that does not impose restrictive i.i.d. assumptions on the input sequence. We have derived conditions for the identifiability of the multichannel system in terms of the underlying channel model orders and the data record length per channel. We have shown that at least 3 channels are required to facilitate separation of the poles and zeros of the system. Simulations highlighting the performance of the technique have been presented.

REFERENCES

[1] L. Tong, G. Xu and T. Kailath, "A new approach to blind identification and equalization of multipath channels," Proc. 25th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, November 1991.

[2] H. Liu, G. Xu, L. Tong, T. Kailath, "Recent Developments in Blind Channel Equalization: From Cyclostationarity to Subspaces," Signal Processing, vol. 50, pp. 83-99, 1996.

[3] T. Kailath, *Linear Systems*, Prentice-Hall, 1980.

[4] A. Gorokhov and P. Loubaton, "Multiple Input Multiple Output ARMA Sytems : Second Order Blind Identification for Signal Extraction," Proc. IEEE SSAP Workshop, pp. 348-351, Corfu, Greece, 1996.

[5] G. Xu, H. Liu, L. Tong and T. Kailath, "A Least-Squares Approach to Blind Channel Identification," IEEE Trans. on Signal Processing, vol. 43, no. 12, December 1995.

[6] C. Scruby, "An Introduction to Acoustic Emission," J. Phys. E. Sci. Instrum., 20:946-953, 1987.

[7] G. T. Venkatesan, D. West, K. M. Buckley, A. H. Tewfik, M. Kaveh, "Automatic Fault Monitoring using Acoustic Emissions," Proc. ICASSP, 1997, Atlanta.

[8] G. T. Venkatesan, M. Kaveh, A. H.Tewfik, K. M. Buckley, "Blind Identification of Single-Input Multiple-Output Pole-Zero Systems," to be submitted, IEEE Transactions on Signal Processing.