A DETERMINISTIC BLIND IDENTIFICATION TECHNIQUE FOR SIMO SYSTEMS OF UNKNOWN MODEL ORDER *

Gopal T. Venkatesan[1] *Lang Tong*[3] *Mos Kaveh*[1] *Ahmed H. Tewfik*[1] *Kevin M. 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.
[3] School of Electrical Engineering,
Cornell University, Ithaca, NY 14853.

ABSTRACT

In this paper we present a method for the deterministic blind identification of single-input multiple-output systems with unknown model order. The technique, that is applicable to both the FIR and IIR cases, requires only an upper bound of the model order. It is based on the special kernel structure of block Toeplitz matrices. When the model order is overestimated, this special structure entails the true solution to be embedded in the overestimated solution in a unique shift-chain form. This special shift-chain structure is then utilized to extract the true solution.

1 INTRODUCTION

Blind identification of single-input multiple-output (SIMO) systems using second-order statistics only, has received widespread interest (see [1][7]) since its feasibility was first demonstrated in [8]. These methods can be widely classified as deterministic (ex. [13]) or stochastic (ex. [2]) according to the model they assume for the source or input sequence. Deterministic blind identification methods possess the finite-sample convergence (FSC) property that provides *exact* estimation with a finite set of data samples, in the absence of noise. Stochastic methods do not exhibit this property. The FSC property is important in situations when the channel is changing rapidly and can be assumed to be stationary only over a short period of time, or when only short records of data are available for the estimation. However, deterministic methods usually require exact prior knowledge of the model order. In situations when this information is not available an estimate of the model order needs to be computed first from the available data. The solutions provided by deterministic methods have been found to be extremely sensitive to the model order. Hence, the use of the model order estimated from a short record of data is unreliable. In contrast, stochastic method require, if at all, only an upper bound of the model order. Though the exact model order might be unknown (or time-varying) in a practical setting, an upper bound on the model order might be known or readily obtained. For example, in a wireless channel, where the dominant cause of intersymbol interference is multipath, the maximum multipath time delay for a variety of scenarios can be estimated by measurements. This would then provide an upper bound on the FIR channel model order. An additional disadvantage associated with deterministic methods is that they are usually formulated for block implementation. Stochastic methods, on the other hand, are amenable to efficient adaptive and recursive implementations.

A deterministic method that requires *only* an estimate on the upper bound of the model order was proposed quite recently in [9][11][14]. By exploiting the isomorphic relation between the channel input and output subspaces, it is shown that the channel order and channel impulse response are uniquely determined by a finite least squares smoothing error sequence in the absence of noise. This least squares smoothing (LSS) approach is, moreover, amenable to recursive implementation. Hence, the LSS technique was the first to demonstrate that the major advantages of stochastic methods could be offered without compromising the FSC property inherent in a deterministic formulation.

In this paper, we will present another deterministic technique that combines the major advantages of stochastic and deterministic methods and requires only an estimate of the upper bound of the model order. Specifically we show that when the model order is overestimated under the deterministic least squares (LS) framework proposed for FIR [13] and pole-zero systems [12], the overestimated LS solution still embeds the true solution in a unique shift-chain form. This is a direct outcome of the null space structure of block Toeplitz matrices. Hence, once the overestimated solution has been obtained, the special shift-chain structure can be utilized to extract the true solution. Therefore, the proposed method, like the LSS method, is a two-step method. The second step of the LSS method and the method proposed in this paper are identical and involve the estimation of the channel parameters from a certain structured subspace that contains the true solution in a shift-chain form. The LSS method and the method proposed in this paper are the only two methods that can provide finite-sample convergence with *only* the upper bound of the model order.

This paper is organized as follows. In Section 2 we present the special kernel structure exhibited by Toeplitz and block Toeplitz matrices. In Section 3 we introduce the deterministic LP problem for the multichannel case and show how the extraneous zeros of the multichannel LP filter can be separated out by using the special shift-chain structure of the block Toeplitz matrix kernel. We then formulate the blind identification problem as a deterministic LP problem in Section 4 and show the applicability of the above two results for the case when the model order is unknown. Simulation results based on the proposed method is presented in Section 5.

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

2 KERNEL STRUCTURE OF BLOCK TOEPLITZ MATRICES

In this section we present the special structure exhibited by the kernel of Toeplitz and block Toeplitz matrices. Detailed treatment of the same can be found in [3].

Let $\mathbf{Y}$ be a $N \times M$ $(N > M)$ Toeplitz matrix. The kernel of $\mathbf{Y}$ can be shown to be the linear hull of a shift-chain. This concept can be expressed compactly using polynomials by utilizing the isomorphism between the space of vectors of dimension M and the space of polynomials of order less than or equal to M. It can be shown that the the kernel of $\mathbf{Y}$ $\{\mathbf{v} \in \mathbf{R}^M : \mathbf{Yv} = \mathbf{0}\}$ is spanned by $\{a(z), za(z), ..., z^{M-Q-1}a(z)\}$, where '$z$' is the shift operator and Q is the order of the polynomial $a(z)$ [3]. That is, the kernel is spanned by,

$$\mathbf{E} = \begin{pmatrix} a(0) & 0 & & 0 \\ a(1) & a(0) & & \vdots \\ \vdots & a(1) & & 0 \\ a(Q) & \vdots & \cdots & a(0) \\ 0 & a(Q) & & a(1) \\ \vdots & 0 & & \vdots \\ 0 & \vdots & & a(Q) \end{pmatrix}. \tag{1}$$

This polynomial $a(z)$ is unique up to a scale factor. In this context it should be noted that when the number of columns M of the matrix $\mathbf{Y}$ is chosen to be Q, the nullity is one. Then the kernel is described by a unique vector characterized by $a(z)$. We will refer to this as the true solution.

The above result extends naturally to the case of block Toeplitz matrices [3]. Let $\mathbf{y} = \{\mathbf{y}(0), \mathbf{y}(1), ..., \mathbf{y}(N+M-2)\}$ be an $N+M-1$ point vector sequence where $\mathbf{y}(i)$'s are $1 \times L$ vectors. Let,

$$\mathbf{Y} = \mathcal{T}_M(\mathbf{y}) = \begin{pmatrix} \mathbf{y_{M-1}} & \mathbf{y_{M-2}} & \cdots & \mathbf{y_0} \\ \mathbf{y_M} & \mathbf{y_{M-1}} & \cdots & \mathbf{y_1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{y_{N+M-2}} & \mathbf{y_{N+M-3}} & \cdots & \mathbf{y_{N-1}} \end{pmatrix} \tag{2}$$

be a block Toeplitz matrix. If $N > ML$ it can be shown that now the kernel of $\mathbf{Y}$ $\{\mathbf{v} \in \mathbf{R}^{ML} : \mathbf{Yv} = \mathbf{0}\}$ is spanned by $\{\mathbf{a}(z), z^L\mathbf{a}(z), ..., z^{L(M-Q-1)}\mathbf{a}(z)\}$, where Q is the order of the $L \times 1$ vector polynomial $\mathbf{a}(z)$ [3].

Summarizing, we can say that if the solution of a set of unknowns is governed by a homogeneous Toeplitz or a block Toeplitz system of equations of the form $\mathbf{Ya} = \mathbf{0}$, then, even if the model order is overestimated, the true solution is still contained in the kernel subspace in a unique shift-chain form as given above.

3 IDENTIFYING EXTRANEOUS ZEROS IN A DETERMINISTIC LP FORMULATION

Consider the homogeneous Toeplitz system of equations $\mathbf{Yv} = \mathbf{0}$ discussed in the previous section. Note that this is a linear prediction formulation of an observed set of data. This deterministic LP formulation has been used for estimating the parameters of a class of pole-zero models [5] for the single channel (or the Toeplitz) case. The unknown LP filter coefficients $\mathbf{a}$ embedded in $\mathbf{v}$ provides complete information on the poles. In the method presented in [5] a singular value decomposition followed by an examination of the singular values provided an estimate of the unknown model order. There the estimation of the exact model order was not critical. This is because the extraneous zeros of the linear prediction filter, that is, the extra zeros introduced into the problem due to overestimation of the model order, can be shown to lie approximately uniformly in a circle within the unit circle [6]. This property of the extraneous zeros is independent of the underlying data if the prediction filter coefficients are chosen to have minimum Euclidean length. Using the above property and conjugate reciprocal root locations of forward and backward data matrices the extraneous zeros can be easily eliminated. Note that these properties hold good only for the single channel (Toeplitz) case. For the case of multiple channels (or the block Toeplitz case), there exists no relation between the root locations for the forward and backward data matrices [4]. Also, the extraneous zeros are not guaranteed to fall within the unit circle. Hence, the separation of the extraneous zeros of the LP filter from the true zeros in the multichannel case does not appear to be straightforward. However, the kernel structure of block Toeplitz matrices as described in the previous section provides a simple alternative procedure.

Consider a $N \times MQ$ block Toeplitz data matrix $\mathbf{Y}$ representing a deterministic LP formulation of an M-channel system such that $\mathbf{Ya} = \mathbf{0}$. The unknown $MQ \times 1$ coefficient vector $\mathbf{a}$ is the true solution and lies in the null space of $\mathbf{Y}$. Note that Q is the true model order and the matrix $\mathbf{Y}$ is of nullity one. Now let $R > Q$ be the overestimated order of the system. The new block Toeplitz matrix $\tilde{\mathbf{Y}}$ of size $(N - R + Q) \times MR$ generated using the same data is of nullity $R - Q + 1$. The kernel structure of $\tilde{\mathbf{Y}}$ is a shift-chain containing the true solution $\mathbf{a}$. Assume now that the true model order is unknown, i.e., Q is unknown. Let

$$\mathbf{T} = \begin{pmatrix} \mathbf{D} & | & \mathbf{E} \end{pmatrix} \tag{3}$$

where $\mathbf{D}$ and $\mathbf{E}$ are matrices that contain a basis for the row and null spaces of $\tilde{\mathbf{Y}}$ respectively. After $\mathbf{T}$ is computed from the data matrix $\tilde{\mathbf{Y}}$ using singular value decomposition, $\mathbf{D}$ and $\mathbf{E}$ may be obtained from $\mathbf{T}$ by suitable thresholding of the singular values. This has been found to be unreliable at medium and lower level SNRs. But noting that the kernel structure is identical to the error subspace obtained via a data projection operation in the LSS approach, we can use the second step outlined therein to overcome the above difficulty. The LSS technique uses a minimum equation error criterion to reliably estimate the nullity [9][11].

Let $\mathbf{d_i}$ be the columns of $\mathbf{D}$. This implies that $\mathbf{d_i}^T\mathbf{E} = 0$. Due to the structure of $\mathbf{E}$, as given in Section 2, we can write this as a set of linear equations in terms of the unknown $\mathbf{a}_q$ for a given order q as [9][11],

$$\begin{pmatrix} \mathbf{D_0} \\ \mathbf{D_1} \\ \cdots \\ \mathbf{D_{MR-(R-q)}} \end{pmatrix} \mathbf{a}_q \tag{4}$$

where $\mathbf{D_i} = \mathcal{T}_q(\mathbf{d_i})$. It is straightforward to see that in the noisefree case the equation error in (4) is zero only when the estimated order is equal to the true order of the system, i.e.,

$q = Q$ [1]. This is because when the model order is underestimated there are insufficient degrees of freedom to specify all the roots defining the null space. When the model order is overestimated, vectors that lie in the null space are incorporated into the row space and leads to an inconsistency in the equations in (4) (see [10]). In the case of perturbation in the estimated right singular space due to additive noise in the observations, the above set of equations can be solved in the LS sense. Hence, the computation of the equation error for various model orders and picking the model order that yields the minimum equation error provides a mechanism for estimating the true model order of the system. These estimates are usually more reliable than estimates obtained directly from the singular values at lower SNRs especially when the system response has small head and tail coefficients [11].

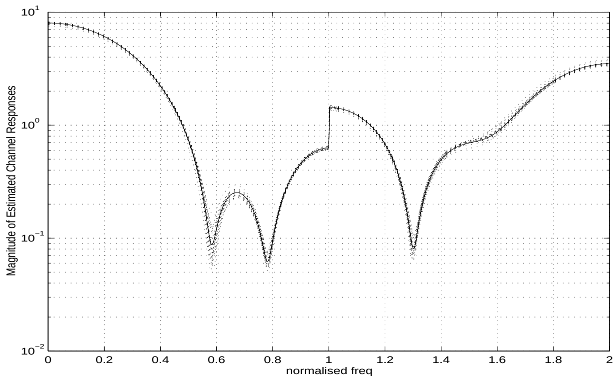

Figure 1. Magnitude of Estimated Channel Responses - Proposed Method

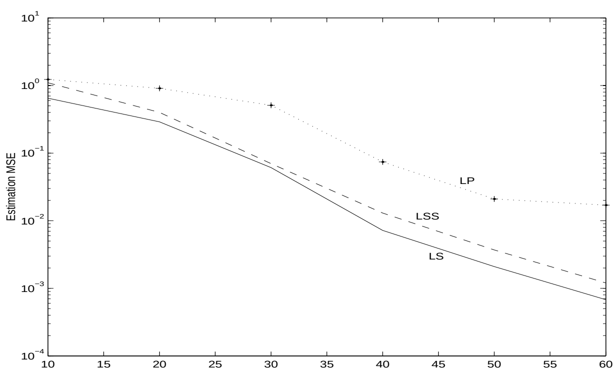

Figure 2. MSE in estimating input sequence

4 DETERMINISTIC LP FORMULATION OF THE BLIND IDENTIFICATION PROBLEM

It was shown in Section 2 that the solution to the overestimated block Toeplitz model contains the true solution in a unique shift-chain fashion. In Section 3 we provided a minimum equation error criterion to identify the extraneous zeros in an overestimated deterministic LP formulation for the multichannel case using the special kernel structure presented in Section 2. Since the LS blind identification technique for both the FIR and pole-zero cases involves the solution of a block Toeplitz system of equations and is equivalent to a deterministic LP formulation, the results of the previous two sections can be applied to combat the problem of unknown model order. Hence, in cases when the model order is not known *a priori* and the model is overestimated to prevent any loss of information, the kernel structure property of block Toeplitz matrices can be used to "discard the extraneous zeros" and obtain the true solution.

We will illustrate the procedure using a two-channel FIR system for simplicity. The procedure can be extended in a straightforward way to the general case of M channels. Let $\mathbf{x_0}$ and $\mathbf{x_1}$ be the output data collected at each channel. Then the LS solution for the unknown channel responses $\mathbf{b_0}$ and $\mathbf{b_1}$ of model order Q is given by [13],

$$\begin{pmatrix} \mathcal{T}_Q(\mathbf{x_1}) & -\mathcal{T}_Q(\mathbf{x_0}) \end{pmatrix} \begin{pmatrix} \mathbf{b_0} \\ \mathbf{b_1} \end{pmatrix} = \mathbf{0} \qquad (5)$$

where $\mathcal{T}_Q(.)$ is the scalar version of the block Toeplitz form defined in (2). In general, when only an estimate of the upper bound of the model order $\tilde{Q}$ is available, the coefficient matrix for the equation set is given by $\mathbf{X} = \begin{pmatrix} \mathcal{T}_{\tilde{Q}}(\mathbf{x_1}) & -\mathcal{T}_{\tilde{Q}}(\mathbf{x_0}) \end{pmatrix}$. Note that this matrix can be transformed to block Toeplitz form by suitable rearrangement of the columns. This results in a corresponding reordering of the elements of the kernel vectors. Using the results in Section 2 it is straightforward to show that the kernel of the above $\mathbf{X}$ matrix is given by,

$$\left(\begin{array}{cccc} b_0(0) & 0 & & 0 \\ b_0(1) & b_0(0) & & \vdots \\ \vdots & b_0(1) & & 0 \\ b_0(Q-1) & \vdots & \cdots & b_0(0) \\ 0 & b_0(Q-1) & & b_0(1) \\ \vdots & 0 & & \vdots \\ 0 & \vdots & & b_0(Q-1) \\ \hline b_1(0) & 0 & & 0 \\ b_1(1) & b_1(0) & & \vdots \\ \vdots & b_1(1) & & 0 \\ b_1(Q-1) & \vdots & \cdots & b_1(0) \\ 0 & b_1(Q-1) & & b_1(1) \\ \vdots & 0 & & \vdots \\ 0 & \vdots & & b_1(Q-1) \end{array} \right) \qquad (6)$$

where $\mathbf{b_0}$ and $\mathbf{b_1}$ are the solutions for the two channels. Hence, the kernel vectors can be used to extract the true solution. Now we mentioned earlier that small deviations in model order have been known to lead to large deviations in the LS solution. This is because, until now the minimum eigenvector of the matrix $\mathbf{X}^T\mathbf{X}$ has been accepted as the LS solution. It is obvious from (6) that this solution is not the true solution *unless* $\tilde{Q} = Q$. Once, the right singular

[1]When this is true, the solution $\mathbf{a_q}$ is equal to $\mathbf{a}$, the true solution of the system.

space of the overestimated model order data matrix $\mathbf{X} = \left(\mathcal{T}_{\tilde{Q}}(\mathbf{x_1}) \quad -\mathcal{T}_{\tilde{Q}}(\mathbf{x_0}) \right)$ has been computed, the minimum equation error criterion method outlined in Section 3 can be used to solve for the true channel parameters.

We summarize the procedure as follows :

Step 1: Choose $\tilde{Q} > Q$, where Q is the unknown true model of the system.

Step 2: Form the data matrix $\mathbf{X}$ using the overestimated order $\tilde{Q}$ as given in [13][12].

Step 3: Compute the SVD of $\mathbf{X}$ to obtain an orthogonal basis for the right singular space.

Step 4: Solve for the true channel parameters using the minimum equation criterion procedure outlined in Section 3.

Once the channel parameters have been estimated, the input sequence can be computed by equalizing the channel. The two-step procedure outlined above can be extended in a straightforward way to any number of channels. Note that a deterministic method proposed recently in [12] for pole-zero systems uses a block Toeplitz set of equations. Hence, the proposed procedure can also be applied to the pole-zero case.

The efficient implementation of the proposed method needs to be studied. But we wish to note here that given the Toeplitz structure of the matrix, the singular space can be updated recursively with every incoming data sample. This is followed by the second step of the LSS method which can also be efficiently implemented [10][11].

5 SIMULATIONS

Consider a two-channel SIMO FIR system. The true model order (unknown *a priori*) was $Q = 5$ and $N = 100$ observation points were available per channel. An upper bound for the model order $R = 9$ was used. Figure 1 shows the superimposed channel frequency response estimates for 20 realizations of additive noise at an SNR of 20 dB. The frequency responses of the two channels are concatenated and plotted. The x-axis 0 to 1 corresponds to the response of the first channel and 1 to 2 corresponds to the response of the second channel. Note that the estimates are quite reliable.

Now we compare the mean squared error (MSE) in estimating the input sequence using the LS technique proposed in this paper and the LS smoothing (LSS) technique. The MSE at various SNRs is plotted in Figure 2. The LS technique does slightly better than the LSS approach for this specific situation. The MSE in estimation for a stochastic method, the LP method [2], is also shown. Note that the MSE "floors-off" at higher SNRs. This is because the LP method being a stochastic method, does not exhibit the FSC property.

6 CONCLUSIONS

In this paper we have presented a technique for the deterministic blind identification of SIMO systems with unknown model order. Under an equivalent multichannel deterministic LP formulation we have shown that the problem of finding the true solution when the model order is overestimated becomes equivalent to discarding the extraneous zeros of the estimated multichannel LP filter. This is accomplished by utilizing the special shift-chain structure of the kernel of block Toeplitz matrices. The two-step method involves, first, the estimation of the right singular space of a block Toeplitz matrix followed by a joint estimation of the true solution and the model order. The second step is identical to that of the LS smoothing approach. The proposed procedure can be applied to both FIR and IIR systems.

ACKNOWLEDGEMENTS

The authors wish to thank Ms. Qing Zhao, Cornell University, for her detailed reading and review of the paper.

REFERENCES

[1] K. Abed-Meraim, W. Qiu and Y. Hua, "Blind system identification," *Proc. IEEE*, vol. 85, no. 8, pp. 1310-1322, August 1997.

[2] K. Abed-Meraim, E. Moulines and P. Loubaton, "Prediction error method for second-order blind identification," *IEEE Trans. Signal Proc.*, vol. 45, no. 3, pp. 694-705, March 1997.

[3] G. Heinig and P. Jankowski, "Kernel structure of block Hankel and Toeplitz matrices and partial realization," *Lin. Algebra and its Applns.*, vol. 175, pp. 1-30, 1992.

[4] S. Kay, *Modern Spectrum Estimation : Theory and Application*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[5] R. Kumaresan and D. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise," *IEEE Trans. on Acous., Speech and Sig. Proc.*, vol. 30, no. 6, pp. 833-840, December 1982.

[6] R. Kumaresan, "On the zeros of the linear prediction-error filter for deterministic signals," *IEEE Trans. on Acous., Speech and Sig. Proc.*, vol. 31, no. 1, pp. 217-220, February 1983.

[7] 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.

[8] L. Tong, G. Xu and T. Kailath, "A new approach to blind identification and equalization of multipath channels," *Proc. 25th Asilomar Conf. on Signals, Sys. and Comp.*, November 1991.

[9] L. Tong and Q. Zhao, "Blind channel estimation by least squares smoothing," *Proc. ICASSP*, Seattle, May 1998.

[10] L. Tong and Q. Zhao, "Joint order detection and blind channel estimation by least squares smoothing," *submitted to IEEE Trans. on Signal Proc.*, Jan. 1998.

[11] L. Tong and Q. Zhao, "Joint order detection and blind channel estimation by least squares smoothing," *Proc. Conf. on Inform. Sci. and Systems*, March 1998.

[12] G. T. Venkatesan, M. Kaveh, A. H. Tewfik and K. M. Buckley, "Blind Identification of Single-Input Multiple-Output Pole-Zero Systems," *Proc. ICASSP*, Seattle, 1998.

[13] G. Xu, H. Liu, L. Tong and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Proc.*, vol. 43, no. 12, pp. 2982-2993, Dec. 1995.

[14] Q. Zhao and L. Tong, "Fast convergence adaptive blind channel estimation and deconvolution," *Proc. Conf. on Inform. Sci. and Systems*, March 1998.