AIAA Meeting Papers on Disc, 1996, pp. 1024-1034
A9626903, AIAA Paper 96-1435

An adaptive Toeplitz/ERA time-domain identification algorithm

James C. Akers
Michigan Univ., Ann Arbor

Dennis S. Bernstein
Michigan Univ., Ann Arbor

IN:AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference and Exhibit, 37th, Salt Lake City, UT, Apr. 15-17, 1996, Technical Papers. Pt. 2 (A96-26801 06-39), Reston, VA, American Institute of Aeronautics and Astronautics, 1996, p. 1024-1034

Two recursive Toeplitz algorithms are used to estimate the Markov parameters of an LTI system from measurements of the inputs and outputs of a system and in turn use ERA to construct a minimal realization. The algorithms can be used either on-line or in an off-line batch mode. The recursive Toeplitz algorithms are shown to be stable and under the assumption of a persistent excitation the estimated Markov parameters converge to the actual Markov parameters. A numerical example of a second-order SISO lightly damped system illustrates the stability and convergence properties of both algorithms. Finally, the algorithms are used to obtain a 20th-order realization of the dynamics of an acoustic duct. (Author)

96-1435

A96-26903 AIAA-96-1435-CP

An Adaptive Toeplitz/ERA Time-Domain Identification Algorithm

by

James C. Akers [1] and Dennis S. Bernstein
Department of Aerospace Engineering, University of Michigan
Ann Arbor, Michigan 48109-2118

Abstract

Two recursive Toeplitz algorithms are used to estimate the Markov parameters of an LTI system from measurements of the inputs and outputs of a system and in turn use ERA to construct a minimal realization. The algorithms can be used either on-line or in an off-line batch mode. The recursive Toeplitz algorithms are shown to be stable and under the assumption of a persistent excitation the estimated Markov parameters converge to the actual Markov parameters. A numerical example of a second-order single-input single-output lightly damped system illustrates the stability and convergence properties of both algorithms. Finally, the algorithms are used to obtain a 20th-order realization of the dynamics of an acoustic duct.

1 Introduction

Real-time identification of the dynamics of linear time-varying systems is becoming a requirement in areas such as health monitoring and damage detection in aircraft and spacecraft structures as well as noise suppression in the interiors of aircraft and automobiles. Traditionally, frequency-domain identification methods are conducted off-line requiring fast Fourier transforms of measured time histories to construct the frequency response function (FRF) of the system. Since these algorithms do not lend themselves to on-line implementation the need exists to develop on-line time-domain identification algorithms. Similarly, the eigensystem realization algorithm (ERA) [3] uses Markov parameters to obtain a minimal realization of the system. However, these Markov parameters are usually obtained from an inverse of the FRF.

Recursive time-domain identification techniques have been studied in the context of neural networks and learning processes. Hyland [4] introduced the concept of a constrained gradient descent approach in identification of finite impulse response systems which maintained the causal structure but not the block-Toeplitz structure of the neural network. We will refer to this as a *pseudo-Toeplitz* type algorithm. Ahn [2] used a batch ARMA model and the update laws of [4] to identify the transfer function coefficients of infinite impulse response systems. Hyland and others [5], [6] introduced the ARMARKOV model structure and gave noise rejection properties of neural networks incorporating this structure.

This paper introduces two recursive time-domain identification algorithms for linear systems which estimate the Markov parameters directly from time-domain data of the system and in turn use ERA to construct a minimal realization. The first algorithm, which we will refer to as the *recursive Toeplitz identification algorithm*, is a generalization of [2]. Vectors comprised of inputs and outputs of the system are used recursively to estimate a weight matrix containing a desired number of Markov parameters. The differences between the recursive Toeplitz algorithm and [2] are an ARMARKOV model structure is used and the estimated weight matrix is constrained to be both causal and block-Toeplitz. The second algorithm, which we will refer to as the *recursive pseudo-Toeplitz identification algorithm*, is also a generalization of [2]. While this algorithm uses an ARMARKOV model structure, the weight matrix is constrained only to be causal resulting in a computationally simpler expression for the gradient. Once the estimated weight matrix of either algorithm has converged, the estimated Markov parameters are extracted and used to construct a Markov block Hankel matrix which is used within ERA to obtain a minimal realization. Both algorithms are well suited to be used either in an on-line or off-line batch mode.

Section 2 introduces ARMARKOV models of linear discrete-time systems which include autoregressive moving-average ARMA models as a special case. Section 3 introduces the recursive Toeplitz identification algorithm and its stability and con-

[1] Member AIAA.

Research supported in part by the Air Force Office of Scientific Research under grants F49620-95-1-0019 and F49620-94-1-0409.

vergence properties. Section 4 introduces the recursive pseudo-Toeplitz identification algorithm and its stability and convergence properties. Section 5 shows how to construct minimal realizations from the Markov block Hankel matrix of which ERA is a special case. In Section 6 both the recursive pseudo-Toeplitz and the recursive Toeplitz identification algorithms are used to identify a second-order single-input single-output (SISO) lightly damped system. In Section 7 the recursive pseudo-Toeplitz/ERA algorithm is used to identify the acoustic dynamics of a duct. Concluding remarks are given in Section 8.

2 ARMARKOV Models

Consider the discrete-time finite-dimensional linear time-invariant system

$$x(k+1) = Ax(k) + Bu(k), \quad (1)$$
$$y(k) = Cx(k) + Du(k), \quad (2)$$

where $A \in \mathcal{R}^{n\times n}$, $B \in \mathcal{R}^{n\times m}$, $C \in \mathcal{R}^{l\times n}$, $D \in \mathcal{R}^{l\times m}$, and n is the order of the realization. The *Markov parameters* H_k are defined by

$$H_k \triangleq D, \quad k=-1, \quad (3)$$
$$\triangleq CA^kB, \quad k \geq 0,$$

which satisfy

$$G(z) = C(zI-A)^{-1}B + D = \sum_{k=-1}^{\infty} H_k z^{-(k+1)}. \quad (4)$$

For convenience let 0_l and $0_{l\times m}$ denote the $l\times l$ and $l\times m$ zero matrices, respectively, and let I_l and $1_{l\times m}$ denote the $l\times l$ identity matrix and $l\times m$ ones matrix, respectively. The transfer function $G(z)$ can also be expressed as

$$G(z) = \frac{B_0z^n + B_1z^{n-1} + \cdots + B_n}{z^n + a_1z^{n-1} + \cdots + a_n}, \quad (5)$$

where $\det(zI-A) = z^n + a_1z^{n-1} + \cdots + a_n$ and $B_i \in \mathcal{R}^{l\times m}$, $i = 0,1,\ldots,n$. Equating (4) and (5) and multiplying both sides by $z^n + a_1z^{n-1} + \cdots + a_n$ yields

$$\begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_n \end{bmatrix} = \begin{bmatrix} H_{-1} & 0 & \cdots & 0 \\ H_0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ H_{n-1} & \cdots & H_0 & H_{-1} \end{bmatrix} \begin{bmatrix} I_m \\ a_1I_m \\ \vdots \\ a_nI_m \end{bmatrix} \quad (6)$$

and

$$H_{n+k} = -\sum_{j=1}^{n} a_j H_{n-j+k}, \quad k \geq 0. \quad (7)$$

Note that the ARMA representation of (5) is given by

$$y(k) = -a_1y(k-1) - \cdots - a_ny(k-n) + B_0u(k) + \cdots + B_nu(k-n), \quad k \geq 0. \quad (8)$$

For a positive integer p, define the *output vector* $Y(k) \in \mathcal{R}^{pl}$ and the *ARMA regressor vector* $\Phi_1(k) \in \mathcal{R}^{(p+n-1)(l+m)+m}$ by

$$Y(k) \triangleq \begin{bmatrix} y(k) \\ \vdots \\ y(k-p+1) \end{bmatrix}, \quad \Phi_1(k) \triangleq \begin{bmatrix} y(k-1) \\ \vdots \\ y(k-p-n+1) \\ u(k) \\ \vdots \\ u(k-p-n+1) \end{bmatrix}. \quad (9)$$

Then the *ARMA/Toeplitz representation* [4] of (5) is given by

$$Y(k) = W_1\Phi_1(k), \quad (10)$$

where the *ARMA weight matrix* W_1 is the block-Toeplitz matrix defined by

$$W_1 \triangleq \left[\begin{matrix} -\mathcal{A}_1 & 0_l & \cdots & 0_l \\ 0_l & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_l \\ 0_l & \cdots & 0_l & -\mathcal{A}_1 \end{matrix} \quad \begin{matrix} H_{-1} & \mathcal{B}_1 & 0_{l\times m} & \cdots & 0_{l\times m} \\ 0_{l\times m} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0_{l\times m} \\ 0_{l\times m} & \cdots & 0_{l\times m} & H_{-1} & \mathcal{B}_1 \end{matrix}\right], \quad (11)$$

where

$$\mathcal{A}_1 \triangleq [a_1I_l \ \cdots \ a_nI_l] \in \mathcal{R}^{l\times nl},$$
$$\mathcal{B}_1 \triangleq [B_1 \ \cdots \ B_n] \in \mathcal{R}^{l\times nm}.$$

Note that W_1 explicitly involves one Markov parameter, namely, H_{-1}.

Next, as introduced in [5], we express $y(k)$ in terms of past inputs, past outputs, and Markov parameters. Substituting (8) with k replaced by $k-1$ back into (8) yields

$$\begin{aligned} y(k) = {} & (a_1^2 - a_2)y(k-2) + (a_1a_2 - a_3)y(k-3) \\ & + \cdots + (a_1a_{n-1} - a_n)y(k-n) \\ & + a_1a_ny(k-n-1) + B_0u(k) \\ & + (B_1 - a_1B_0)u(k-1) \\ & + (B_2 - a_1B_1)u(k-2) \\ & + \cdots + (B_n - a_1B_{n-1})u(k-n) \\ & - a_1B_nu(k-n-1). \end{aligned} \quad (12)$$

Noting from (6) that

$$H_{-1} = B_0, \quad H_0 = B_1 - a_1B_0, \quad (13)$$

and substituting (13) into (12) yields

$$\begin{aligned} y(k) = {} & \alpha_{2,1}y(k-2) + \cdots + \alpha_{2,n}y(k-n-1) \\ & + H_{-1}u(k) + H_0u(k-1) \\ & + \mathcal{B}_1u(k-2) + \cdots + \mathcal{B}_nu(k-n-1), \end{aligned} \quad (14)$$

where

$$\alpha_{2,i} \triangleq a_1 a_i - a_{i+1}\ ,\quad i = 1,\ldots,n-1,$$
$$\alpha_{2,n} \triangleq a_1 a_n,$$
$$\mathcal{B}_{2,i} \triangleq B_{i+1} - a_1 B_i\ ,\quad i = 1,\ldots,n-1,$$
$$\mathcal{B}_{2,n} \triangleq a_1 B_n.$$

Since (14) explicitly involves the first two Markov parameters of $G(z)$, it is called an *ARMARKOV representation.* Defining the *ARMARKOV regressor vector* $\Phi_2(k) \in \mathcal{R}^{(p+n-1)(l+m)+2m}$ by

$$\Phi_2(k) \triangleq \begin{bmatrix} y(k-2) \\ \vdots \\ y(k-p-n) \\ u(k) \\ \vdots \\ u(k-p-n) \end{bmatrix}, \qquad (15)$$

yields the *ARMARKOV/Toeplitz representation* [5] of (5) given by

$$Y(k) = W_2 \Phi_2(k), \qquad (16)$$

where the *ARMARKOV weight matrix* W_2 is the block-Toeplitz matrix is defined by

$$W_2 \triangleq \left[\begin{array}{cccccccccc} -\mathcal{A}_2 & 0_l & \cdots & 0_l & H_{-1} & H_0 & \mathcal{B}_1 & 0_{l\times m} & \cdots & 0_{l\times m} \\ 0_l & \ddots & \ddots & \vdots & 0_{l\times m} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_l & \vdots & \ddots & \ddots & \ddots & \ddots & 0_{l\times m} \\ 0_l & \cdots & 0_l & -\mathcal{A}_2 & 0_{l\times m} & \cdots & 0_{l\times m} & H_{-1} & H_0 & \mathcal{B}_2 \end{array}\right], \qquad (17)$$

where

$$\mathcal{A}_2 \triangleq [\alpha_{2,1} I_l \ \cdots \ \alpha_{2,n} I_l] \in \mathcal{R}^{l\times nl},$$
$$\mathcal{B}_2 \triangleq [\mathcal{B}_{2,1} \ \cdots \ \mathcal{B}_{2,n}] \in \mathcal{R}^{l\times nm}.$$

Note that W_2 involves two Markov parameters, namely, H_{-1} and H_0.

An ARMARKOV model whose weight matrix contains the first three Markov parameters can be obtained by substituting $y(k-2)$ given by (8) into (14). Repeating this procedure yields an ARMARKOV model possessing the first μ Markov parameters of the form

$$\begin{aligned} y(k) &= \sum_{j=1}^{n} -\alpha_{\mu,j} y(k-\mu-j+1) \\ &\quad + \sum_{j=1}^{\mu} H_{j-2} u(k-j+1) \\ &\quad + \sum_{j=1}^{n} \mathcal{B}_{\mu,j} u(k-\mu-j+1). \end{aligned} \qquad (18)$$

Defining the *ARMARKOV regressor vector* $\Phi_\mu(k) \in \mathcal{R}^{(p+n-1)(l+m)+\mu m}$ by

$$\Phi_\mu(k) \triangleq \begin{bmatrix} y(k-\mu) \\ \vdots \\ y(k-\mu-p-n+2) \\ u(k) \\ \vdots \\ u(k-\mu-p-n+2) \end{bmatrix}, \qquad (19)$$

it follows that

$$Y(k) = W_\mu \Phi_\mu(k), \qquad (20)$$

where the *ARMARKOV weight matrix* W_μ is the block-Toeplitz matrix defined by

$$W_\mu \triangleq \left[\begin{array}{cccccccccc} -\mathcal{A}_\mu & 0_l & \cdots & 0_l & H_{-1} & \cdots & H_{\mu-2} & \mathcal{B}_\mu & 0_{l\times m} & \cdots & 0_{l\times m} \\ 0_l & \ddots & \ddots & \vdots & 0_{l\times m} & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_l & \vdots & \ddots & \ddots & & \ddots & \ddots & 0_{l\times m} \\ 0_l & \cdots & 0_l & -\mathcal{A}_\mu & 0_{l\times m} & \cdots & 0_{l\times m} & H_{-1} & \cdots & H_{\mu-2} & \mathcal{B}_\mu \end{array}\right] \qquad (21)$$

where

$$\mathcal{A}_\mu \triangleq [\alpha_{\mu,1} I_l \ \cdots \ \alpha_{\mu,n} I_l] \in \mathcal{R}^{l\times nl},$$
$$\mathcal{B}_\mu \triangleq [\mathcal{B}_{\mu,1} \ \cdots \ \mathcal{B}_{\mu,n}] \in \mathcal{R}^{l\times nm}.$$

Note that the ARMA model is a specialized ARMARKOV model with $\mu = 1$, while setting $\mu = 1$ and $\mu = 2$ in (21) yields (11) and (17).

Remark 2.1 Note that the transfer function of (18) can be expressed as

$$G(z) = \frac{H_{-1} z^{\mu+n-1} + \cdots + H_{\mu-2} z^n + \mathcal{B}_1 z^{n-1} + \cdots + \mathcal{B}_n}{z^{\mu+n-1} + \alpha_1 z^{n-1} + \cdots + \alpha_n}.$$

If $G(z)$ is asymptotically stable then $\lim_{\mu\to\infty} \alpha_{\mu,i} = 0$, $\lim_{\mu\to\infty} \mathcal{B}_{\mu,i} = 0$, $i = 1,\ldots,n$. Therefore, if $G(z)$ is stable then as μ increases the ARMARKOV model becomes less sensitive to errors in the $\alpha_{\mu,i}$ and $\mathcal{B}_{\mu,i}$ terms.

Henceforth, for convenience we drop the subscript μ and write W and $\Phi(k)$ for W_μ and $\Phi_\mu(k)$.

3 Recursive Toeplitz Algorithm

Let $\widehat{W}(k)$ denote the *estimated ARMARKOV weight matrix* at time k which is constructed to have the same structure as W, let $\widehat{Y}(k)$ denote the *estimated output vector* defined by

$$\widehat{Y}(k) \triangleq \widehat{W}(k)\Phi(k) \in \mathcal{R}^{pl}, \qquad (22)$$

and define the *output error* $\varepsilon(k)$ by

$$\varepsilon(k) \triangleq Y(k) - \widehat{Y}(k). \tag{23}$$

The *error function* $J(k)$ is defined by

$$J(k) \triangleq \frac{1}{2}\varepsilon^{\mathrm{T}}(k)\varepsilon(k). \tag{24}$$

The recursive Toeplitz algorithm constrains the gradient of $J(k)$ with respect to $\widehat{W}(k)$, and hence $\widehat{W}(k)$, to have the same block-Toeplitz structure as W. Therefore, the independent variables in $\widehat{W}(k)$ are $\widehat{\alpha}_1(k),\ldots,\widehat{\alpha}_n(k)$, $\widehat{H}_1(k),\ldots,\widehat{H}_{\mu-2}(k)$, and $\widehat{\beta}_1(k),\ldots,\widehat{\beta}_n(k)$. Let

$$\begin{aligned}
\xi_i(k) &\triangleq \widehat{\alpha}_i(k), \quad i = 1\ldots n,\\
&\triangleq \widehat{H}_{i+\mu-n-3}(k), \quad i = n+1,\ldots,n+\mu,\\
&\triangleq \widehat{\beta}_{i-n-\mu}(k), \quad i = n+\mu+1,\ldots,2n+\mu.
\end{aligned} \tag{25}$$

Furthermore, for convenience we define the following partitioning matrices for constructing the gradient of $J(k)$:

$$\begin{aligned}
L_{i,j} &\triangleq \begin{bmatrix} -1 \\ 0_{(pl-1)\times 1}\end{bmatrix}, \quad i = 1,\ldots,n, \quad j = 1,\\
&\triangleq \begin{bmatrix} 0_{(j-1)l\times 1} \\ -1 \\ 0_{(pl-j)\times 1}\end{bmatrix}, \quad i = 1,\ldots,n,\\
&\qquad j = 2,\ldots,pl-1,\\
&\triangleq \begin{bmatrix} 0_{(pl-1)\times 1} \\ -1 \end{bmatrix}, \quad i = 1,\ldots,n, \quad j = pl,\\
&\triangleq \begin{bmatrix} I_l \\ 0_{(p-1)l\times l}\end{bmatrix}, \quad i = n+1,\ldots,2n+\mu,\\
&\qquad j = 1,\\
&\triangleq \begin{bmatrix} 0_{(j-1)l\times l} \\ I_l \\ 0_{(p-j)l\times l}\end{bmatrix}, \quad i = n+1,\ldots,2n+\mu,\\
&\qquad j = 2,\ldots,p-1,\\
&\triangleq \begin{bmatrix} 0_{(p-1)l\times l} \\ I_l \end{bmatrix}, \quad i = n+1,\ldots,2n+\mu,\\
&\qquad j = p,\\
R_{i,j} &\triangleq [1 \quad 0_{1\times(q-1)}], \; i = 1, \; j = 1,\\
&\triangleq [0_{1\times((i-1)+j-1)} \quad 1 \quad 0_{1\times(q-i-j+1)}],\\
&\qquad i = 1,\ldots,n, \quad j = 1,\ldots,pl, \quad i+j > 2,\\
&\triangleq [0_{m\times(p+i+j-3)m} \quad I_m\\
&\qquad 0_{m\times(2n+p+\mu-i-j)m}],\\
&\qquad i = n+1,\ldots,2n+\mu, \quad j = 1,\ldots,p\\
&\qquad i+j < 2n+\mu+p,\\
&\triangleq [0_{m\times(q-m)} \quad I_m], \quad i = 2n+\mu, \quad j = p.
\end{aligned}$$

Then $\widehat{W}(k)$ can be written as

$$\begin{aligned}
\widehat{W}(k) &= \sum_{i=1}^{n}\sum_{j=1}^{pl} L_{i,j}\widehat{\xi}_i(k)R_{i,j}\\
&\quad + \sum_{i=n+1}^{2n+\mu}\sum_{j=1}^{p} L_{i,j}\widehat{\xi}_i(k)R_{i,j}.
\end{aligned} \tag{26}$$

Lemma 3.1 Consider the error function (24) with the independent variables (25). Then

$$\begin{aligned}
\frac{\partial J(k)}{\partial \widehat{\xi}_i(k)} &\triangleq -\sum_{j=1}^{pl} L_{i,j}^{\mathrm{T}}\varepsilon(k)\Phi^{\mathrm{T}}(k)R_{i,j}^{\mathrm{T}}, \quad i = 1,\ldots,n,\\
&\triangleq -\sum_{j=1}^{p} L_{i,j}^{\mathrm{T}}\varepsilon(k)\Phi^{\mathrm{T}}(k)R_{i,j}^{\mathrm{T}},\\
&\qquad i = n+1,\ldots,2n+\mu.
\end{aligned} \tag{27}$$

Furthermore,

$$\begin{aligned}
\frac{\partial J(k)}{\partial \widehat{W}(k)} &\triangleq \sum_{i=1}^{n}\sum_{j=1}^{pl} L_{i,j}\frac{\partial J(k)}{\partial \widehat{\xi}_i(k)}R_{i,j}\\
&\quad + \sum_{i=n+1}^{2n+\mu}\sum_{j=1}^{p} L_{i,j}\frac{\partial J(k)}{\partial \widehat{\xi}_i(k)}R_{i,j}.
\end{aligned} \tag{28}$$

Note that if $\frac{\partial J(k)}{\partial \widehat{W}(k)} \neq 0$ then $\varepsilon(k) \neq 0$ and $\Phi(k) \neq 0$.

We now consider the *Toeplitz/ARMARKOV weight matrix update law*

$$\widehat{W}(k+1) = \widehat{W}(k) - \eta(k)\frac{\partial J(k)}{\partial \widehat{W}(k)}, \tag{29}$$

where $\eta(k)$ is the *adaptive step size.* Finally, define the *ARMARKOV weight matrix error* by

$$E(k) \triangleq W - \widehat{W}(k). \tag{30}$$

Then it follows from (30) and (29) that

$$E(k+1) = E(k) + \eta(k)\frac{\partial J(k)}{\partial \widehat{W}(k)}, \tag{31}$$

and

$$\varepsilon(k) = E(k)\Phi(k). \tag{32}$$

Theorem 3.1 Consider the update law (29), assume that $\frac{\partial J(k)}{\partial \widehat{W}(k)} \neq 0$, $k \geq 0$, and suppose the adaptive step size $\eta(k)$ satisfies

$$0 < \eta(k) < \frac{2p\|\varepsilon(k)\|_2^2}{\left\|\frac{\partial J(k)}{\partial \widehat{W}(k)}\right\|_{\mathrm{F}}^2}. \tag{33}$$

Then $\{\|E(k)\|_{\mathrm{F}}\}_{k=0}^{\infty}$ is decreasing. Furthermore, if $\eta(k) = \eta_{\mathrm{opt}}(k)$, where

$$\eta_{\mathrm{opt}}(k) \triangleq \frac{p\|\varepsilon(k)\|_2^2}{\left\|\frac{\partial J(k)}{\partial \widehat{W}(k)}\right\|_{\mathrm{F}}^2}, \tag{34}$$

then $\|E(k+1)\|_{\mathrm{F}}^2 - \|E(k)\|_{\mathrm{F}}^2$ is minimized. If, in addition,

$$\sup_{k\geq 0}\left|\frac{\eta(k)}{\eta_{\mathrm{opt}}(k)} - 1\right| < 1, \tag{35}$$

and

$$0 < \inf_{k\geq 0}\|\Phi(k)\|_2 \leq \sup_{k\geq 0}\|\Phi(k)\|_2 < \infty, \tag{36}$$

then
$$\sum_{k=0}^{\infty} \|\varepsilon(k)\|_2^2 < \infty. \quad (37)$$
Consequently,
$$\lim_{k\to\infty} \varepsilon(k) = 0, \quad (38)$$
and
$$\lim_{k\to\infty} \frac{\partial J(k)}{\partial \widehat{W}(k)} = 0. \quad (39)$$

Remark 3.1 If we chose the adaptive step size
$$\eta(k) = \alpha\eta_{\rm opt}(k)\ ,\ \ 0 < \alpha < 2, \quad (40)$$
or the computationally simpler adaptive step size of
$$\eta(k) = \frac{1}{pl^2\|\Phi(k)\|_2^2}\ , \quad (41)$$
then (35) holds.

Remark 3.2 If $\frac{\partial J(k)}{\widehat{W}(k)} = 0$ then it follows from update law (29) that $\widehat{W}(k+1) = \widehat{W}(k)$ and hence $\|E(k+1)\|_{\rm F} = \|E(k)\|_{\rm F}$. Therefore, it follows from Theorem 3.1 that the update law (29) with the the adaptive step size (33) is stable is the sense that $\{\|E(k)\|_{\rm F}\}_{k=0}^{\infty}$ is nonincreasing.

Thus far we have shown that the update law (29) with either the adaptive step size (40) or (41) causes the output error vector $\varepsilon(k)$ to converge to zero if the inputs and outputs are bounded from below away from zero and bounded from above. However, this condition on the input and output sequences is not strong enough to guarantee that $\widehat{W}(k)$ converges to W. To guarantee $\widehat{W}(k)$ converges to W we require a persistent excitation if using the adaptive step size (41) or a strongly persistent excitation if using the adaptive step size (40). Let $\sigma_{\rm min}(\cdot)$ denote the minimum singular value. For convenience let $q_1 = (p+n-1)l$, $q_2 = (p+n-1)m + \mu m$, and $q = (p+n-1)(l+m) + \mu m$ so that $\Phi(k) \in \mathcal{R}^q$.

Definition 3.1 The sequence $\{u(k)\}_{k=0}^{\infty}$ is *persistent* with respect to the ARMARKOV/Toeplitz representation of $G(z)$ if the following three conditions are satisfied:

i) $\{u(k)\}_{k=0}^{\infty}$ and $\{y(k)\}_{k=0}^{\infty}$ are bounded.

ii) $$\inf_{k\geq 0} \left\| \begin{bmatrix} u(k) \\ \vdots \\ u(k-\mu-p-n+2) \end{bmatrix} \right\|_2 > 0\ .$$

iii) There exists $\delta > 0$ such that for all $k > 0$ there exist $k \leq r_1(k) < \cdots < r_q(k)$ such that $\Phi(r_1(k)), \ldots, \Phi(r_q(k))$ are linearly independent and $\sigma_{\rm min}\left[\Phi(r_1(k)) \ \ \ldots \ \ \Phi(r_q(k))\right] > \delta$.

Definition 3.2 The sequence $\{u(k)\}_{k=0}^{\infty}$ is *strongly persistent* with respect to the ARMARKOV/Toeplitz representation of $G(z)$ if the following three conditions are satisfied:

i) $\{u(k)\}_{k=0}^{\infty}$ and $\{y(k)\}_{k=0}^{\infty}$ are bounded.

ii) $$\inf_{k\geq 0} \sigma_{\rm min} \left(\begin{bmatrix} u^{\rm T}(k) \\ \vdots \\ u^{\rm T}(k-p+1) \end{bmatrix} \right) > 0.$$

iii) There exists $\delta > 0$ such that for all $k > 0$ there exist $k \leq r_1(k) < \cdots < r_q(k)$ such that $\Phi(r_1(k)), \ldots, \Phi(r_q(k))$ are linearly independent and $\sigma_{\rm min}\left[\Phi(r_1(k)) \ \ \ldots \ \ \Phi(r_q(k))\right] > \delta$.

Lemma 3.2 If the sequence $\{u(k)\}_{k=0}^{\infty}$ is strongly persistent then it is also persistent.

Note that $G(z)$ is asymptotically stable if and only if it is bounded-input bounded-output stable. Hence, if $G(z)$ is asymptotically stable and $\{u(k)\}_{k=0}^{\infty}$ is bounded then $\{y(k)\}_{k=0}^{\infty}$ is bounded. However, we do not assume that $G(z)$ is asymptotically stable since a bounded $\{u(k)\}_{k=0}^{\infty}$ can be constructed which produces a bounded $\{y(k)\}_{k=0}^{\infty}$ even if $G(z)$ is unstable.

Theorem 3.2 Consider the update law (29). Assume $\{u(k)\}_{k=0}^{\infty}$ is persistent with respect to $G(z)$ and let the adaptive step size $\eta(k)$ be given by (41). Then
$$\lim_{k\to\infty} \widehat{W}(k) = W. \quad (42)$$
Moreover, if $\{u(k)\}_{k=0}^{\infty}$ is strongly persistent with respect to $G(z)$ and the adaptive step size $\eta(k)$ is given by (40) then (42) holds.

Next we consider the special case of a finite impulse response filter and show that if the input sequence $\{u(k)\}_{k=0}^{\infty}$ is an impulse then $\widehat{W}(k)$ converges to W in a finite number of steps.

Lemma 3.3 Consider the SISO finite impulse response filter
$$G(z) = H_{-1} + H_0 z^{-1} + \cdots + H_{r-2} z^{-r+1}, \quad (43)$$
with the update law (29) and the adaptive step size (40). Furthermore, let
$$\begin{aligned} u(k) &= 1\ ,\ \ k = 0, \\ &= 0\ ,\ \ k \geq 1, \end{aligned}$$
and $\mu \geq r$. Then
$$\widehat{W}(k) = W\ ,\ \ k \geq \mu. \quad (44)$$
If, in addition, $\widehat{W}(0) = 0$, then
$$\widehat{W}(k) = W\ ,\ \ k \geq r. \quad (45)$$

Theorem 3.2 shows that if a persistent or strongly persistent excitation is used, depending upon the adaptive step size chosen, then $\widehat{W}(k)$ converges to W and therefore, we can obtain the the Markov parameters $H_{-1}, \ldots, H_{\mu-2}$.

4 Recursive Pseudo-Toeplitz Algorithm

For convenience let $\mathcal{I}_{l\times n} = [I_l \ \cdots \ I_l] \in \mathcal{R}^{l\times nl}$ and "∘" denote the Hadamard product of two matrices. We now constrain $\widehat{W}(k)$ to have the same block-zero or causal structure as W. Furthermore, we assume the remaining entries of $\widehat{W}(k)$ are independent, which does not explicitly take into account the block-Toeplitz structure of W. Hence, the gradient of $J(k)$ in now defined to be

$$\frac{\partial J(k)}{\widehat{W}(k)} \triangleq -U \circ \left(\varepsilon(k)\Phi^{\mathrm{T}}(k)\right), \tag{46}$$

where $U \in \mathcal{R}^{pl\times((p+n-1)(l+m)+\mu m)}$ is defined by

$$U \triangleq \left[\begin{array}{ccccc|cccc} \mathcal{I}_{l\times n} & 0_l & \cdots & 0_l & 1_{l\times(n+\mu-1)m} & 0_{l\times m} & \cdots & 0_{l\times m} \\ 0_l & \ddots & \ddots & \vdots & 0_{l\times m} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_l & \vdots & \ddots & \ddots & 0_{l\times m} \\ 0_l & \cdots & 0_l & \mathcal{I}_{l\times n} & 0_{l\times m} & \cdots & 0_{l\times m} & 1_{l\times(n+\mu-1)m} \end{array}\right]. \tag{47}$$

Note that $\frac{\partial J(k)}{\widehat{W}(k)}$ is not necessarily a block-Toeplitz matrix.

Lemma 4.1 $\frac{\partial J(k)}{\widehat{W}(k)}\Phi(k) = 0$ if and only if $\frac{\partial J(k)}{\widehat{W}(k)} = 0$.

We now consider the ARMARKOV *weight matrix update law*

$$\widehat{W}(k+1) = \widehat{W}(k) - \eta(k)\frac{\partial J(k)}{\widehat{W}(k)}, \tag{48}$$

where $\eta(k)$ is the *adaptive step size*, and define the *predicted output error* $\widehat{\varepsilon}(k)$ by

$$\widehat{\varepsilon}(k) \triangleq Y(k) - \widehat{W}(k+1)\Phi(k), \tag{49}$$

and the *predicted error function* $\widehat{J}(k)$ defined by

$$\widehat{J}(k) \triangleq \frac{1}{2}\widehat{\varepsilon}^{\mathrm{T}}(k)\widehat{\varepsilon}(k). \tag{50}$$

Note that if $\frac{\partial J(k)}{\widehat{W}(k)} = 0$ then $\widehat{W}(k+1) = \widehat{W}(k)$, $\widehat{\varepsilon}(k) = \varepsilon(k)$, and $\widehat{J}(k)$ is independent of $\eta(k)$. Let $\|\cdot\|_2$ and $\|\cdot\|_{\mathrm{F}}$ denote the spectral and Frobenius norms, respectively.

Lemma 4.2 Consider the predicted error function (50) with ARMARKOV weight matrix update law (48) and the predicted output error (49). Then

$$\widehat{J}(k) = \frac{1}{2}\left\|\frac{\partial J(k)}{\partial\widehat{W}(k)}\Phi(k)\right\|_2^2 \eta^2(k) - \left\|\frac{\partial J(k)}{\partial\widehat{W}(k)}\right\|_{\mathrm{F}}^2 \eta(k) + \frac{1}{2}\|\varepsilon(k)\|_2^2. \tag{51}$$

Proposition 4.1 If $\frac{\partial J(k)}{\widehat{W}(k)}\Phi(k) \neq 0$, then the adaptive step size $\eta(k) = \eta_{\mathrm{opt}\widehat{J}(k)}(k)$ defined by

$$\eta_{\mathrm{opt}\widehat{J}(k)}(k) \triangleq \frac{\left\|\frac{\partial J(k)}{\partial\widehat{W}(k)}\right\|_{\mathrm{F}}^2}{\left\|\frac{\partial J(k)}{\partial\widehat{W}(k)}\Phi(k)\right\|_2^2}, \tag{52}$$

is positive and minimizes $\widehat{J}(k)$. Furthermore, if $\eta(k) = \eta_{\mathrm{opt}\widehat{J}(k)}(k)$ then

$$\widehat{J}(k) = -\frac{1}{2}\frac{\left\|\frac{\partial J(k)}{\partial\widehat{W}(k)}\right\|_{\mathrm{F}}^4}{\left\|\frac{\partial J(k)}{\partial\widehat{W}(k)}\Phi(k)\right\|_2^2} + \frac{1}{2}\|\varepsilon(k)\|_2^2. \tag{53}$$

The following result considers the adaptive step size proposed in [2].

Proposition 4.2 Let $G(z)$ be an $l \times m$ transfer function. The adaptive step size $\eta(k) = \eta_{\mathrm{opt}}(k)$ defined by

$$\eta_{\mathrm{opt}}(k) \triangleq \frac{\|\varepsilon(k)\|_2^2}{\left\|\frac{\partial J(k)}{\partial\widehat{W}(k)}\right\|_{\mathrm{F}}^2}, \tag{54}$$

satisfies

$$\eta_{\mathrm{opt}\widehat{J}(k)}(k) = \eta_{\mathrm{opt}}(k), \quad p = 1 \text{ and } l = 1, \tag{55}$$

$$\leq \eta_{\mathrm{opt}}(k), \quad p \geq 1 \text{ or } l \geq 1, \tag{56}$$

for all $k \geq 0$. Furthermore, if $\eta(k) = \eta_{\mathrm{opt}}(k)$ then

$$\widehat{J}(k) = \frac{1}{2}\frac{\left\|\frac{\partial J(k)}{\partial\widehat{W}(k)}\Phi(k)\right\|_2^2}{\left\|\frac{\partial J(k)}{\partial\widehat{W}(k)}\right\|_{\mathrm{F}}^4}\|\varepsilon(k)\|_2^4 - \frac{1}{2}\|\varepsilon(k)\|_2^2, \tag{57}$$

and

$$\varepsilon^{\mathrm{T}}(k)\widehat{\varepsilon}(k) = 0. \tag{58}$$

If, in addition, $p = 1$ and $l = 1$ then

$$\widehat{\varepsilon}(k) = 0. \tag{59}$$

Finally, it follows from the update law (48) that

$$E(k+1) = E(k) + \eta(k)\frac{\partial J(k)}{\widehat{W}(k)}. \tag{60}$$

Theorem4.1 will show that $\eta_{\mathrm{opt}}(k)$ to be optimal with respect to minimizing $\|E(k+1)\|_{\mathrm{F}}^2 - \|E(k)\|_{\mathrm{F}}^2$.

Theorem 4.1 Consider the update law (48), assume that $\frac{\partial J(k)}{\widehat{W}(k)} \neq 0$, $k \geq 0$, and assume that the adaptive step size $\eta(k)$ satisfies

$$0 < \eta(k) < 2\eta_{\rm opt}(k). \qquad (61)$$

Then $\{\|E(k)\|_{\rm F}\}_{k=0}^{\infty}$ is decreasing. Furthermore, if $\eta(k) = \eta_{\rm opt}(k)$ then $\|E(k+1)\|_{\rm F}^2 - \|E(k)\|_{\rm F}^2$ is minimized. If, in addition,

$$\sup_{k\geq 0}\left|\frac{\eta(k)}{\eta_{\rm opt}(k)} - 1\right| < 1, \qquad (62)$$

and

$$0 < \inf_{k\geq 0}\|\Phi(k)\|_2 \leq \sup_{k\geq 0}\|\Phi(k)\|_2 < \infty, \qquad (63)$$

then

$$\sum_{k=0}^{\infty}\|\varepsilon(k)\|_2^2 < \infty. \qquad (64)$$

Consequently,

$$\lim_{k\to\infty}\varepsilon(k) = 0, \qquad (65)$$

and

$$\lim_{k\to\infty}\frac{\partial J(k)}{\widehat{W}(k)} = 0. \qquad (66)$$

Remark 4.1 If we chose the adaptive step size $\eta(k)$ such that

$$\eta(k) = \alpha\eta_{\rm opt}(k)\ , \quad 0 < \alpha < 2, \qquad (67)$$

then (62) holds.

Remark 4.2 If $\frac{\partial J(k)}{\widehat{W}(k)} = 0$ then it follows from update law (48) that $\widehat{W}(k+1) = \widehat{W}(k)$ and hence $\|E(k+1)\|_{\rm F} = \|E(k)\|_{\rm F}$. Therefore, it follows from Theorem 4.1 that the update law (48) with the the adaptive step size (61) is stable is the sense that $\{\|E(k)\|_{\rm F}\}_{k=0}^{\infty}$ is nonincreasing.

Thus far we have shown that the update law (48) with either the adaptive step size (52) or (67) causes the output error vector $\varepsilon(k)$ to converge to zero if the inputs and outputs are bounded from below away from zero and bounded from above. However, this condition on the input and output sequences is not strong enough to guarantee that $\widehat{W}(k)$ converges to W. To guarantee $\widehat{W}(k)$ converges to W requires a persistent excitation.

Theorem 4.2 Consider the update law with the adaptive step size given by (52) or (67). If $\{u(k)\}_{k=0}^{\infty}$ is persistent with respect to $G(z)$ then

$$\lim_{k\to\infty}\widehat{W}(k) = W. \qquad (68)$$

Next we consider the special case of a finite impulse response filter and show that if the input sequence $\{u(k)\}_{k=0}^{\infty}$ is an impulse then $\widehat{W}(k)$ converges to W in a finite number of steps.

Lemma 4.3 Consider the SISO finite impulse response filter

$$G(z) = H_{-1} + H_0 z^{-1} + \cdots + H_{r-2}z^{-r+1}, \qquad (69)$$

with the update law (48) with either the adaptive step size (52) or (54). Furthermore, let

$$\begin{aligned} u(k) &= 1\ , \quad k = 0, \\ &= 0\ , \quad k \geq 1, \end{aligned}$$

and $\mu \geq r$. Then

$$\widehat{W}(k) = W\ , \quad k \geq \mu + p - 1. \qquad (70)$$

If, in addition, $\widehat{W}(0) = 0$, then

$$\widehat{W}(k) = W\ , \quad k \geq r + p - 1. \qquad (71)$$

Theorem 4.2 shows that if a persistent excitation is used then $\widehat{W}(k)$ converges to W and therefore, we can obtain the the Markov parameters $H_{-1}, \ldots, H_{\mu-2}$. In the next section we use these Markov parameters to construct state-space realizations.

5 Minimal Realizations from Markov Parameters

This section gives a brief overview of [1] showing how minimal realizations are constructed using different decompositions of the the Markov block Hankel matrix $\mathcal{H}_{r,s}(0)$. Let $\widehat{H}_j(k)$ denote the estimation of the Markov parameter H_j obtained from the AR-MARKOV weight matrix $\widehat{W}(k)$. For positive integers r and s and for $k \geq -1$ the *Markov block Hankel matrix* $\mathcal{H}_{r,s,j} \in \mathcal{R}^{(r+1)l\times(s+1)m}$ is defined by

$$\mathcal{H}_{r,s,j} \triangleq \begin{bmatrix} H_j & \cdots & H_{j+s} \\ \vdots & \ddots & \vdots \\ H_{j+r} & \cdots & H_{j+r+s} \end{bmatrix}. \qquad (72)$$

Let $G(z)$ denote the transfer function whose AR-MARKOV weight matrix is W. We begin by stating a well-known result concerning the rank of $\mathcal{H}_{r,s,0}$ [7, p. 442], which in turn specifies the minimum number of Markov parameters necessary for constructing a minimal realization.

Lemma 5.1 Assume $G(z)$ has McMillan degree n, and let $r, s \geq n-1$. Then rank $\mathcal{H}_{r,s,0} = n$.

For convenience let $E_{i,j} \triangleq \begin{bmatrix} I_j \\ 0_{ij \times j} \end{bmatrix}$. Let $G(z) \sim \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ denote the transfer function corresponding to the state-space realization (1) and (2). The notation « $\overset{\min}{\sim}$ » denotes a minimal realization. The following result provides a method for constructing a realization of $G(z)$ from a sufficiently large but finite number of Markov parameters.

Theorem 5.1 Let $G(z)$ have McMillan degree n, let $r, s \geq n-1$, and let $P \in \mathcal{R}^{(r+1)l \times n}$, $\Sigma_{r,s} \in \mathcal{R}^{n \times n}$, and $Q \in \mathcal{R}^{n \times (s+1)m}$ be such that $\Sigma_{r,s}$ is positive definite and

$$\mathcal{H}_{r,s,0} = P \Sigma_{r,s} Q. \tag{73}$$

Furthermore, let $P_{\rm L}$ and $Q_{\rm R}$ denote left and right inverses of P and Q, respectively. Then

$$G(z) \overset{\min}{\sim} \left[\begin{array}{c|c} \Sigma_{r,s}^{-1/2} P_{\rm L} \mathcal{H}_{r,s,1} Q_{\rm R} \Sigma_{r,s}^{-1/2} & \Sigma_{r,s}^{1/2} Q E_{s,m} \\ \hline E_{r,l}^{\rm T} P \Sigma_{r,s}^{1/2} & H_{-1} \end{array}\right]. \tag{74}$$

Moreover, the s-stage controllability and r-stage observability Gramians of (74) are given by

$$W_{\mathcal{C}_s} = \Sigma_{r,s}^{1/2} Q Q^{\rm T} \Sigma_{r,s}^{1/2}, \quad W_{\mathcal{O}_r} = \Sigma_{r,s}^{1/2} P^{\rm T} P \Sigma_{r,s}^{1/2}. \tag{75}$$

Next we choose the factorization (73) according to the Singular Value Decomposition. In this case Theorem 5.1 yields the Eigensystem Realization Algorithm (ERA) [3].

Corollary 5.1 Let $G(z)$ have McMillan degree n, and let $r, s \geq n-1$. Furthermore, let $P \in \mathcal{R}^{(r+1)l \times n}$, $\Sigma_{r,s} \in \mathcal{R}^{n \times n}$, and $Q \in \mathcal{R}^{n \times (s+1)m}$ satisfy (73), where $P^{\rm T}P = QQ^{\rm T} = I$ and $\Sigma_{r,s} = \operatorname{diag}(\sigma_1^{r,s}, \ldots, \sigma_n^{r,s})$, where $\sigma_1^{r,s} \geq \cdots \geq \sigma_n^{r,s} > 0$ are the singular values of $\mathcal{H}_{r,s,0}$. Then

$$G(z) \overset{\min}{\sim} \left[\begin{array}{c|c} \Sigma_{r,s}^{-1/2} P^{\rm T} \mathcal{H}_{r,s,1} Q^{\rm T} \Sigma_{r,s}^{-1/2} & \Sigma_{r,s}^{1/2} Q E_{s,m} \\ \hline E_{r,l}^{\rm T} P \Sigma_{r,s}^{1/2} & H_{-1} \end{array}\right]. \tag{76}$$

Moreover, this realization is (r,s)-finitely balanced with

$$W_{\mathcal{C}_s} = W_{\mathcal{O}_r} = \Sigma_{r,s}. \tag{77}$$

The structure of the ERA realization (76) suggests that if the last $n-q$ entries of Σ are small compared to the first q entries, then truncating the last $n-q$ states of (76) will result in a reduced-order realization that retains the dominant dynamic characteristics of the actual system $G(z)$. Therefore, let $\mathcal{H}_{r,s,0} = P\Sigma_{r,s}Q$ where $P \in \mathcal{R}^{(r+1)l \times n}$, $\Sigma_{r,s} \in \mathcal{R}^{n \times n}$, $Q \in \mathcal{R}^{n \times (s+1)m}$, $P^{\rm T}P = QQ^{\rm T} = I_n$, and $\Sigma_{r,s} = \operatorname{diag}(\sigma_1^{r,s}, \ldots, \sigma_n^{r,s})$. Letting P_1, Σ_1, and Q_1 denote the first q columns, the leading $q \times q$ submatrix, and the first q rows of P, $\Sigma_{r,s}$, and $\widehat{Q}$, respectively, then a reduced-order realization that approximates $G(z)$ is given by

$$\widetilde{G}(z) \sim \left[\begin{array}{c|c} \Sigma_1^{-1/2} P_1^{\rm T} \mathcal{H}_{r,s,1} Q_1^{\rm T} \Sigma_1^{-1/2} & \Sigma_1^{1/2} Q_1 E_{s,m} \\ \hline E_{r,l}^{\rm T} P_1 \Sigma_1^{1/2} & H_{-1} \end{array}\right]. \tag{78}$$

For a system of McMillan degree n ERA requires $r, s \geq n-1$. Note that for $r, s = n-1$ the highest indexed Markov parameter in $\mathcal{H}_{r,s,1}$ is $H_{1+r+s} = H_{2n-1}$. Therefore, ERA requires at least the first $2n+1$ Markov parameters in order to identify a system of McMillan degree n.

6 Numerical Example

In this section both the recursive pseudo-Toeplitz and recursive Toeplitz algorithms are used to obtain estimates of the first six Markov parameters of a second-order SISO discrete-time system. This numerical example is a single degree-of-freedom oscillator with a natural frequency $f_{\rm n}$ of 10 Hz and a damping ratio ρ of 1%. The transfer function of this continuous-time second-order system is given by

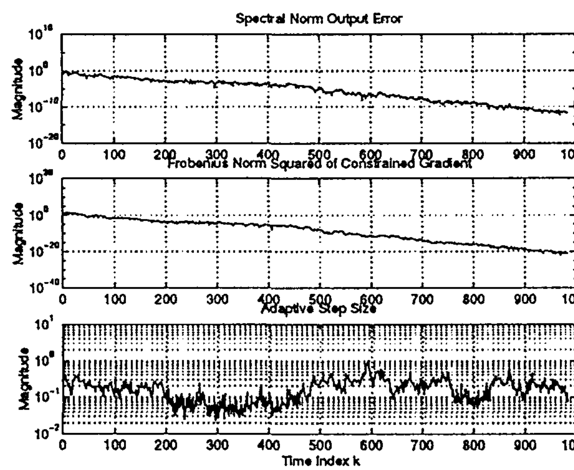

Figure 1: Pseudo-Toeplitz Algorithm, $\mu = 6$, $n = 2$, $p = 4$, White Noise .

$$G(s) = \frac{\omega_{\rm n}^2}{s^2 + 2\rho\omega_{\rm n}s + \omega_{\rm n}^2}, \tag{79}$$

with a minimal realization given by

$$\begin{aligned} \dot{x} &= \begin{bmatrix} 0 & 1 \\ -\omega_{\rm n}^2 & -2\rho\omega_{\rm n} \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \omega_{\rm n} \end{bmatrix} u(t), \\ y &= [\omega_{\rm n} \quad 0] x(t) + [0] u(t). \end{aligned} \tag{80}$$

A zero-order-hold discretization of (80) at a sampling frequency of 100 Hz is given by

$$x(k+1) = \begin{bmatrix} 8.0981e-1 & 9.2964e-3 \\ -3.6701e+1 & 7.9813e-1 \end{bmatrix} x(k), + \begin{bmatrix} 3.0270e-3 \\ 5.8411e-1 \end{bmatrix} u(k)$$
$$y(k) = [6.2832e+1 \quad 0]\, x(k) + [0]\, u(k). \quad (81)$$

6.1 Recursive Pseudo-Toeplitz Algorithm

The convergence of the recursive pseudo-Toeplitz algorithm with the ARMARKOV model having $\mu = 6$, $n = 2$, and $p = 4$ and the input being zero-mean uncorrelated random noise with a uniform distribution and a standard deviation of $\frac{1}{\sqrt{3}}$ is shown in Figures 1 and 2. Note that $q = (p+n-1)(l+m)+\mu m = 16$. The $\|[u(k) \;\cdots\; u(k-10)]\|_2$ and

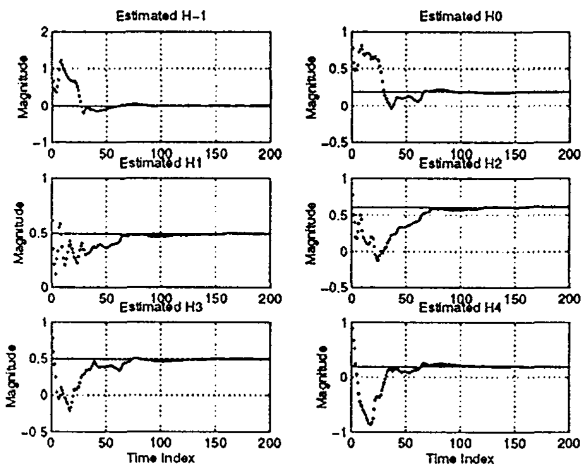

Figure 2: Pseudo-Toeplitz Algorithm, $\mu = 6$, $n = 2$, $p = 4$, Convergence of Estimated Markov Parameters, White Noise.

$\sigma_{\min}([\Phi(k) \;\cdots\; \Phi(k+15)])$ are shown in Figure 3 where $\|[u(k) \;\cdots\; u(k-10)]\|_2 > 0.5$, $k \geq 0$ and $\sigma_{\min}([\Phi(k) \;\cdots\; \Phi(k+15)]) > 2.2e-16$, the machine precision, at a significant number of time steps indicating the input is persistent. The six estimated Markov parameters were obtained by averaging over the corresponding entries of $\widehat{W}(k)$ and are shown to have converged after only 100 steps in Figure 2. Note the optimal adaptive step size $\eta_{\mathrm{opt}}(k)$ remains approximately constant while $\left\|\frac{\partial J(k)}{\partial W(k)}\right\|_F^2$ decreases exponentially in Figure 1. Although not shown, performance of this algorithm with $\eta_{\mathrm{opt}\widehat{J}(k)}$ was indistinguishable to that when using the $\eta_{\mathrm{opt}}(k)$ since $\eta_{\mathrm{opt}\widehat{J}(k)}$ was only slightly smaller that $\eta_{\mathrm{opt}}(k)$.

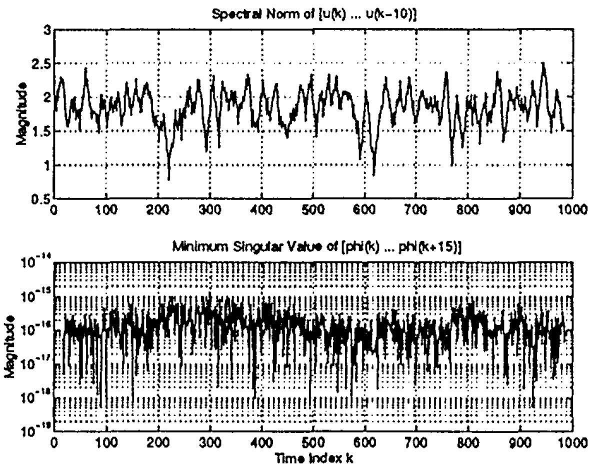

Figure 3: Pseudo-Toeplitz Algorithm, $\mu = 6$, $n = 2$, $p = 4$, White Noise .

6.2 Recursive Toeplitz Algorithm

The convergence of the recursive Toeplitz algorithm with the ARMARKOV model having $\mu = 6$, $n = 2$, and $p = 4$ and the input being zero-mean uncorrelated random noise with a uniform distribution and a standard deviation of $\frac{1}{\sqrt{3}}$ is shown in Figure 4. The six estimated Markov parameters are shown to have converged after approximately 70 time steps.

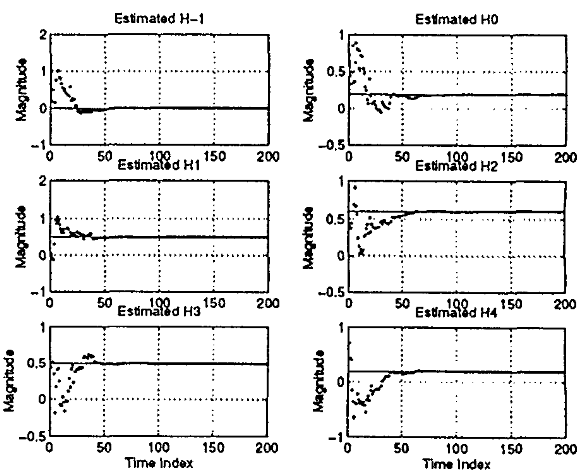

Figure 4: Recursive Toeplitz Algorithm, $\mu = 6$, $n = 2$, $p = 4$, Convergence of Estimated Markov Parameters, White Noise.

7 Identification of an Acoustic Duct

In this section the pseudo-Toeplitz/ERA algorithm was used to identify the transfer function from a speaker amplifier input to a microphone amplifier output from 0 to 400 Hz. A spectrum analyzer with a display window bandwidth of 400 Hz, sampling frequency of 1024 Hz, was used to produce the white-noise excitation signal as well as record both FRF's and the input and output time-histories. The

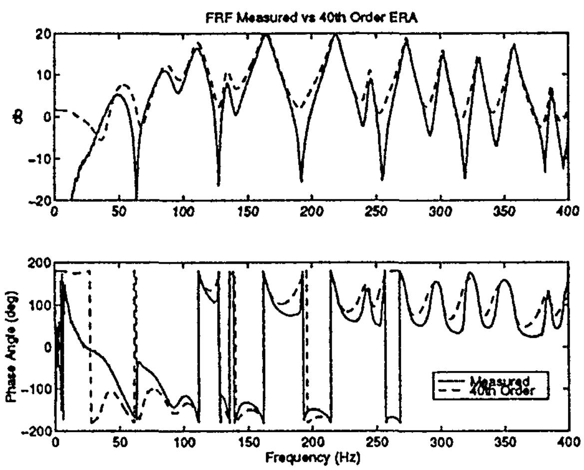

Figure 5: Frequency-Domain ID of Acoustic Duct.

speaker was mounted on the side of an open-closed acoustic duct 19.75 feet long. All of the modes of the duct between 0 and 400 Hz were purely longitudinal modes. The microphone was mounted very close to the speaker compared to the wavelength of the highest frequency mode below 400 Hz and therefore, was considered to be colocated with respect to the speaker. For comparison purposes, the inverse of a measured FRF was used with ERA in order to obtain a frequency-domain identified model. The size of the Markov block Hankel matrices used in ERA was chosen to be 100×100. The frequency-domain analysis produced a 40th-order model which is plotted along with the measured FRF in Figure 5.

The input and output time- histories consisted of 4096 data points with a length of approximately 4 seconds. In order to be able to make a direct comparison to the frequency-domain identified model the Markov Block Hankel matrix was chosen to be 100×100 which required $\mu \geq 201$. The ARMARKOV model was chosen to have $\mu = 210$, $n = 40$, and $p = 4$. The algorithm is shown in Figure 6 to have stopped converging after 3000 data points or approximately

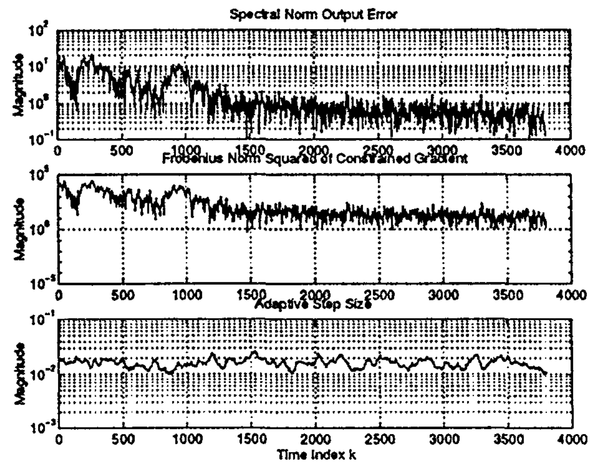

Figure 6: Pseudo-Toeplitz Algorithm, $\mu = 210$, $n = 40$, $p = 4$, Acoustic Duct.

3 seconds of data. The estimated Markov parameters were obtained from the total input and output time- histories. The 20th order model obtained with ERA and shown in Figure 7 was the best. Note that the estimated Markov parameters caused ERA to produced an unstable 40th order model. In fact all models above 20th order were either poor fits or unstable. Finally, Figure 8 shows the convergence of 20th order identified models of the adaptive Toeplitz/ERA algorithm after 1, 2, 3, and 4 seconds of data. While the 40th order frequency-domain

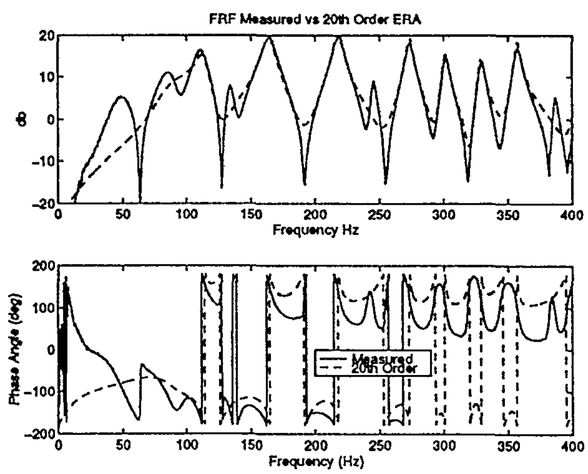

Figure 7: Pseudo-Toeplitz/ERA Time-Domain ID of Acoustic Duct.

model does better approximate the measured FRF of the duct compared to the 20th order time-domain model, the time-domain model still captures the most significant duct modes. Hence, the adaptive Toeplitz/ERA algorithm has provided a reasonable model of the duct.

8 Conclusions

We have introduced two recursive time-domain identification algorithms that estimate the Markov parameters, based upon an ARMARKOV model structure, and in turn use ERA to construct a minimal realization. Under the assumption that the the adaptive step size and the spectral norm of the regressor vector are appropriately bounded it has been shown that both algorithms are stable and converge. If, in addition, the input is persistent or strongly persistent, depending upon the algorithm and adaptive step size chosen, then the estimated ARMARKOV weight matrix converges to the actual ARMARKOV weight matrix. The numerical exam-

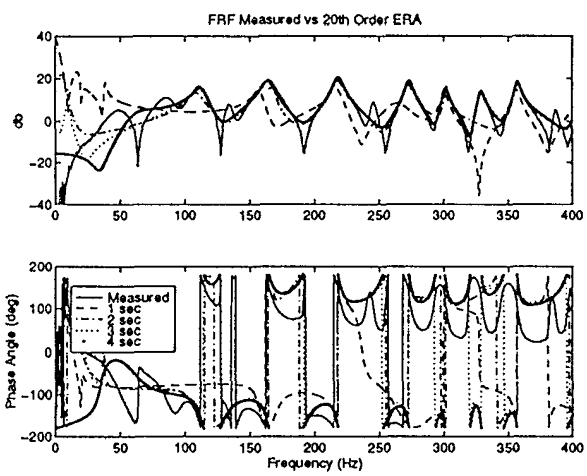

Figure 8: Pseudo-Toeplitz/ERA Time-Domain ID of Acoustic Duct.

ple of a second-order lightly damped SISO system shows convergence of the estimated Markov parameters for a white noise input. Moreover, the recursive Toeplitz algorithm is shown to converge faster than the recursive pseudo-Toeplitz algorithm. However, it should be noted that the recursive Toeplitz algorithm is computationally much more intensive making the recursive pseudo-Toeplitz algorithm better suited to on-line implementation. The effect of the output vector length upon the performance of the algorithm still needs to be explored. We conjecture that as the output vector is lengthened, the noise rejection properties of the algorithm should be enhanced.

Finally, the recursive pseudo-Toeplitz/ERA algorithm was applied off-line successfully to identify the dynamics of an acoustic duct and generate a relatively high order realization.

References

[1] J. C. Akers and D. S. Bernstein, "Measurement Noise Error Bounds for Eigensystem Realization Algorithm," *Proc. 36th AIAA Structures, Structural Dynamics, Materials Conf.*, pp 1198-1208, New Orleans, LO, April, 1995.

[2] C. Ahn, "An recursive Algorithm for Identification of IIR Systems," *Proc. IEEE Conf. Dec. Contr.*, pp. 2595-2600, San Antonio, TX, December, 1993.

[3] J. Juang and R. Pappa, "An Eigensystem Realization algorithm for Modal Parameter Identification and Model Reduction," *Journal of Guidance, Control, and Dynamics*, Vol. 8, no. 5, pp 620-627, 1985.

[4] D. C. Hyland, "Neural Network Architecture for On-Line System Identification and recursively Optimized Control," *Proc. IEEE Conf. Dec. Contr.*, pp. 2552-2557, Brighton, UK, December, 1991.

[5] D. C. Hyland, "Neural Networks System Identification," Final Report to NASA Langley Research Center, Contract NAS1-18225, pp 18-19, October, 1993.

[6] D. C. Hyland, E. G. Collins Jr., W. M. Haddad, and D. L. Hunter, "Neural Network System Identification for Improved Noise Rejection," *Proc. Amer. Contr. Conf.*, pp 345-349, Seattle, WA, June, 1995.

[7] T. Kailath, *Linear Systems*, Prentice-Hall, New Jersey, 1980.