

Adaptive Disturbance Rejection Using ARMARKOV/Toeplitz Models

Ravinder Venugopal and Dennis S. Bernstein, *Member, IEEE*

Abstract—An adaptive disturbance rejection algorithm is developed for the standard control problem. The multiple input–multiple output (MIMO) system and controller are represented as ARMARKOV/Toeplitz models, and the parameter matrix of the compensator is updated on-line by means of a gradient algorithm. The algorithm requires minimal knowledge of the plant, specifically, the numerator of the ARMARKOV model of the transfer function from the control inputs to the performance variables is required. No knowledge about the spectrum of the disturbance is needed. Experimental results demonstrating tonal and broadband disturbance rejection in an acoustic duct are presented.

Index Terms—Active noise and vibration control, adaptive control, discrete-time systems, disturbance rejection.

NOMENCLATURE

$0_{l \times m}$	$l \times m$ zero matrix.
I_l	$l \times l$ identity matrix.
$1_{l \times m}$	$l \times m$ ones matrix.
$\mathcal{I}_{l \times m}$	$[I_l \dots I_l] \in \mathcal{R}^{l \times ml}$.
$\ \cdot\ _2$	Euclidean vector norm.
$\ \cdot\ _F$	Frobenius matrix norm.

I. INTRODUCTION

AN important objective of control system design is to minimize the effects of external disturbance signals. For applications such as active noise and vibration control, it is the primary focus. In cases where the system is time varying or difficult to identify, adaptive methods such as the feedforward least mean square (LMS) and recursive LMS (RLMS) algorithms are useful [1]–[7]. However, feedforward-type algorithms neglect the effect of the feedback path from control to measurement thus leading to poor performance and instability [8]. To remedy this problem, robust variations of the classical LMS algorithm have been proposed; see, for example [9].

Predictive models, which involve the Markov parameters of the system, are used in predictive control of systems with time delays [10, pp. 169–179], [11, pp. 106–110], [12, pp. 331–365], [13]–[16]. Markov-parameter-based representations of systems also provide a framework for direct controller synthesis based

on input–output data [17]. In addition, predictive control algorithms such as the long range generalized predictive algorithm [12, pp. 353–362], [13] use windows of data. Predictive models are also used in [18]–[21] for model identification within recursive and batch least squares techniques. In these works predictive models are termed *ARMARKOV models* to emphasize the presence of Markov parameters in ARMA-type models. In [20] it is shown that ARMARKOV models can be used to estimate Markov parameters in the presence of persistent, but not necessarily white, input signals. In [22] it is shown that least-squares identification yields consistent estimates of the Markov parameters in the presence of persistent measurements.

In the present paper we develop an adaptive disturbance rejection controller using ARMARKOV plant and controller models. Our approach is distinct from predictive control techniques due to the fact that the adaptation mechanism we employ is based upon past data rather than future predicted error. A gradient algorithm that minimizes a retrospective performance cost function is used to update the entries of the controller parameter matrix. The update law uses an adaptive step size involving past data and an ARMARKOV model of the secondary path transfer function.

We begin in Section III by formulating the disturbance rejection problem in terms of the standard two-input two-output (TITO) framework. In Section IV, we review ARMARKOV models of discrete-time finite-dimensional linear time-invariant systems and derive the ARMARKOV model for TITO systems. Next, in Section V, we develop an adaptive algorithm for disturbance rejection by representing the controller in terms of an ARMARKOV parameter matrix and deriving a gradient-descent-based update law for this matrix corresponding to the retrospective performance cost function. The algorithm requires performance measurement sensors which may also be used as feedback measurement sensors. However, additional feedback measurement sensors may be used to enhance closed-loop performance. The algorithm does not require a direct measurement of the disturbance. The Markov parameters and moving average coefficients that relate the performance to the control (the secondary path transfer function) need to be known to implement the algorithm. These parameters are obtained by using the time-domain identification algorithm of [20] and [21]. Since ARMARKOV models are used for identification and since the controller is based upon such a representation, the intermediate step of recovering a state-space or frequency domain model of the system is eliminated. The last section presents experimental results for tonal and broadband disturbance rejection in an acoustic duct.

Manuscript received August 25, 1997; revised August 3, 1998. Recommended by Editor-in-Chief, M. Bodson. This work was supported in part by the Air Force Office of Scientific Research under Grant F49620-98-1-0037.

The authors are with the Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109-2140 USA (e-mail: ds-baero@engin.umich.edu).

Publisher Item Identifier S 1063-6536(00)01787-5.

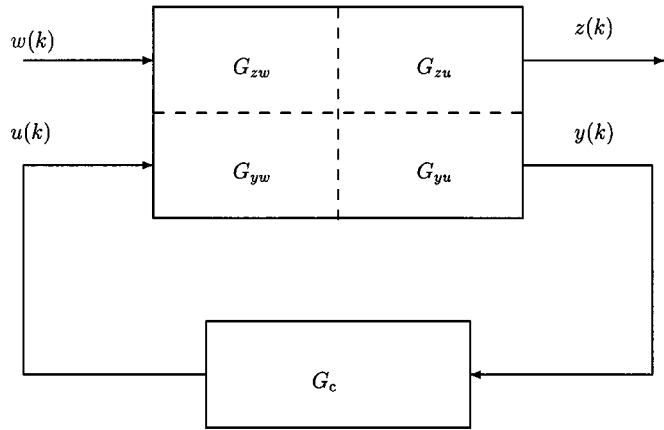


Fig. 1. Standard problem with fixed-gain controller.

II. STANDARD PROBLEM REPRESENTATION OF DISTURBANCE REJECTION

Consider the linear discrete-time TITO system shown in Fig. 1. The disturbance $w(k)$, the control $u(k)$, the measurement $y(k)$ and the performance $z(k)$ are in \mathcal{R}^{m_w} , \mathcal{R}^{m_u} , \mathcal{R}^{l_y} , and \mathcal{R}^{l_z} , respectively. The system can be written in state-space form as

$$x(k+1) = Ax(k) + Bu(k) + D_1w(k) \quad (1)$$

$$z(k) = E_1x(k) + E_2u(k) + E_0w(k) \quad (2)$$

$$y(k) = Cx(k) + Du(k) + D_2w(k) \quad (3)$$

or equivalently in terms of transfer matrices

$$z = G_{zw}w + G_{zu}u \quad (4)$$

$$y = G_{yw}w + G_{yu}u. \quad (5)$$

The controller G_c generates the control signal $u(k)$ based on the measurement $y(k)$, that is,

$$u = G_c y. \quad (6)$$

The objective of the standard problem [25] is to determine a controller G_c that produces a control signal $u(k)$ based on the measurement $y(k)$ such that a performance measure involving $z(k)$ is minimized. In classical fixed-gain H_2 and H_∞ optimal control theory, the performance $z(k)$ is not required to be measured, but rather G_{zw} and G_{yw} are used analytically for off-line controller design (Fig. 1). Fixed-gain controller design methods for disturbance rejection generally require knowledge of all four transfer matrices, namely, the primary path G_{zw} , the secondary path G_{zu} , the reference path G_{yw} and the feedback path G_{yu} , as well as the spectrum of the disturbance $w(k)$. This terminology is standard in the noise control literature [2] and the feedforward structure described therein assumes that $G_{yu} = 0$ and $G_{yw} = I$.

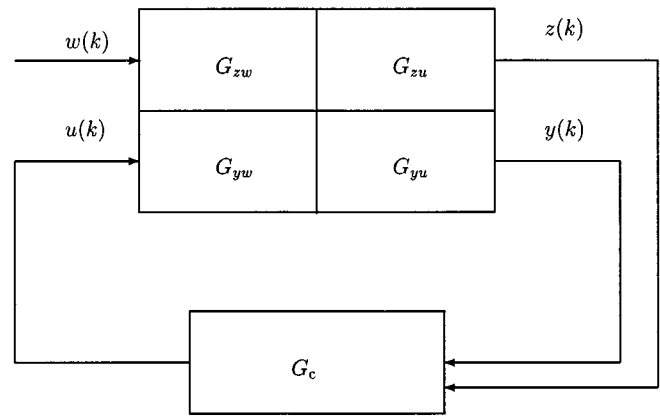


Fig. 2. Standard problem with performance assumption.

Unlike fixed-gain controller design methods, adaptive control techniques require on-line measurement of $z(k)$ for use in adaptation. If $z(k)$ is measured and used for control, we say that the performance assumption is satisfied, as shown in Fig. 2. In the case that $z(k)$ is not available as a measurement, a filter that estimates $z(k)$ based on available signals can be used. However, in contrast to fixed-gain methods, adaptive methods [1]–[6] often require that only the secondary path transfer matrix G_{zu} be known. Other adaptive methods [9] identify G_{zu} on-line but require additional actuators and sensors.

III. ARMARKOV MODELS

In this section we derive the ARMARKOV representation of a state space model. Consider the n th-order discrete-time finite-dimensional linear time-invariant system

$$x(k+1) = Ax(k) + Bu(k) \quad (7)$$

$$y(k) = Cx(k) + Du(k) \quad (8)$$

where $u(k) \in \mathcal{R}^{m_u}$ and $y(k) \in \mathcal{R}^{l_y}$. The Markov parameters $H_j \in \mathcal{R}^{l_y \times m_u}$ of this system are defined as

$$H_j \triangleq D, \quad j = -1 \quad (9)$$

$$\triangleq CA^j B, \quad j \geq 0 \quad (10)$$

and satisfy

$$G(z) \triangleq C(zI - A)^{-1}B + D = \sum_{j=-1}^{\infty} H_j z^{-(j+1)}. \quad (11)$$

Note that the Markov parameters are the impulse response coefficients of the system. The transfer function $G(z)$ can be equivalently represented as

$$G(z) = \frac{1}{z^n + a_1 z^{n-1} + \dots + a_n} \cdot (B_0 z^n + B_1 z^{n-1} + \dots + B_n) \quad (12)$$

where $\det(zI - A) = z^n + a_1z^{n-1} + \dots + a_n$ and $B_i \in \mathcal{R}^{l_y \times m_u}, i = 0, \dots, n$. Equating (11) and (12) and multiplying both sides by $z^n + a_1z^{n-1} + \dots + a_n$ yields

$$\begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_n \end{bmatrix} = \begin{bmatrix} H_{-1} & 0 & \cdots & 0 \\ H_0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ H_{n-1} & \cdots & H_0 & H_{-1} \end{bmatrix} \begin{bmatrix} I_m \\ a_1 I_m \\ \vdots \\ a_n I_m \end{bmatrix} \quad (13)$$

which provides recursive expressions for B_i in terms of H_i . Now consider the ARMA representation of (12) given by

$$y(k) = -a_1y(k-1) - \dots - a_ny(k-n) + B_0u(k) + \dots + B_nu(k-n). \quad (14)$$

Replacing k with $k-1$ in (14) and substituting the resulting relation back into (14) yields

$$\begin{aligned} y(k) &= (a_1^2 - a_2)y(k-2) + (a_1a_2 - a_3)y(k-3) \\ &+ \dots + (a_1a_{n-1} - a_n)y(k-n) \\ &+ a_1a_ny(k-n-1) + B_0u(k) + (B_1 - a_1B_0)u(k-1) \\ &+ (B_2 - a_1B_1)u(k-2) \\ &+ \dots + (B_n - a_1B_{n-1})u(k-n) \\ &- a_1B_nu(k-n-1). \end{aligned} \quad (15)$$

Noting from (13) that $H_{-1} = B_0$, $H_0 = B_1 - a_1B_0$, and defining

$$\begin{aligned} \alpha_{2,i} &\triangleq a_1a_i - a_{i+1}, \quad i = 1, \dots, n-1 \\ \alpha_{2,n} &\triangleq a_1a_n \end{aligned} \quad (16)$$

$$\begin{aligned} \mathcal{B}_{2,i} &\triangleq B_{i+1} - a_1B_i, \quad i = 1, \dots, n-1 \\ \mathcal{B}_{2,n} &\triangleq a_1B_n \end{aligned} \quad (17)$$

(15) can be written as

$$\begin{aligned} y(k) &= \alpha_{2,1}y(k-2) + \dots + \alpha_{2,n}y(k-n-1) \\ &+ H_{-1}u(k) + H_0u(k-1) + \mathcal{B}_{2,1}u(k-2) \\ &+ \dots + \mathcal{B}_{2,n}u(k-n-1). \end{aligned} \quad (18)$$

We note that (18) explicitly involves the first two Markov parameters H_{-1} and H_0 , and thus is called a 2-ARMARKOV

model. Repeating this procedure $\mu - 1$ times yields the μ -ARMARKOV model of (7) and (8)

$$\begin{aligned} y(k) &= \sum_{j=1}^n -\alpha_j y(k - \mu - j + 1) + \sum_{j=1}^{\mu} H_{j-1} u(k - j + 1) \\ &+ \sum_{j=1}^n \mathcal{B}_j u(k - \mu - j + 1) \end{aligned} \quad (19)$$

where $\alpha_j \in \mathcal{R}$ and $\mathcal{B}_j \in \mathcal{R}^{l_y \times m_u}, j = 1, \dots, n$.

Equation (19) is an input-output relation that explicitly involves μ Markov parameters. For $\mu > 1$, (19) has the same form as the μ step ahead predictor [10, pp. 169–179], [12, pp. 136–139], while in the case $\mu = 1$, (19) specializes to the usual ARMA model. The coefficients α_j and \mathcal{B}_j can be calculated recursively using (14) and (13). Equation (19) can be equivalently represented as the ARMARKOV transfer function

$$G(z) = \frac{1}{z^{\mu+n-1} + \alpha_1 z^{n-1} + \dots + \alpha_n} \cdot (H_{-1} z^{\mu+n-1} + \dots + H_{\mu-2} z^n + \mathcal{B}_1 z^{n-1} + \dots + \mathcal{B}_n). \quad (20)$$

This system representation is nonminimal, overparameterized, and constrained since the numerator and denominator of the transfer function are polynomials of order $\mu + n - 1$, and the coefficients of the terms $z^{\mu+n-2}$ through z^n in the denominator are zero.

Now, let p be a positive integer and define the extended measurement vector $Y(k) \in \mathcal{R}^{lp}$ and the ARMARKOV regressor vector $\Phi_{yu}(k) \in \mathcal{R}^{l_y(p+n-1)+m_u(\mu+p+n-1)}$ by

$$\begin{aligned} Y(k) &\triangleq \begin{bmatrix} y(k) \\ \vdots \\ y(k-p+1) \end{bmatrix} \\ \Phi_{yu}(k) &\triangleq \begin{bmatrix} y(k-\mu) \\ \vdots \\ y(k-\mu-p-n+2) \\ u(k) \\ \vdots \\ u(k-\mu-p-n+2) \end{bmatrix}. \end{aligned} \quad (21)$$

Using (19), $Y(k)$ and $\Phi_{yu}(k)$ are related by

$$Y(k) = W_{yu} \Phi_{yu}(k) \quad (22)$$

$W_{yu} \triangleq$

$$\begin{bmatrix} -\alpha_1 I_{l_y} & \cdots & -\alpha_n I_{l_y} & 0_{l_y} & \cdots & 0_{l_y} & H_{-1} & \cdots & H_{\mu-2} & \mathcal{B}_1 & \cdots & \mathcal{B}_n & 0_{l_y \times m_u} & \cdots & 0_{l_y \times m_u} \\ 0_{l_y} & \ddots & \ddots & \ddots & \ddots & \vdots & 0_{l_y \times m_u} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0_{l_y} & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0_{l_y \times m_u} \\ 0_{l_y} & \cdots & 0_{l_y} & -\alpha_1 I_{l_y} & \cdots & -\alpha_n I_{l_y} & 0_{l_y \times m_u} & \cdots & 0_{l_y \times m_u} & H_{-1} & \cdots & H_{\mu-2} & \mathcal{B}_1 & \cdots & \mathcal{B}_n \end{bmatrix}. \quad (23)$$

where the block-Toeplitz *ARMA* weight matrix $W_{yu} \in \mathcal{R}^{p_l y \times [l_y(p+n-1) + m_u(\mu+p+n-1)]}$ is defined by (23) shown at the bottom of the previous page. We note that a state-space realization of the system can be obtained from (19) by either constructing a canonical form or using the eigensystem realization algorithm (ERA) [20], [21].

IV. ARMA/TOEPLITZ MODEL OF TITO SYSTEMS

We now develop the ARMA/Toeplitz model of (1)–(3). Defining the Markov parameters of the system by

$$H_{yu,-1} \triangleq D, \quad H_{yu,j} \triangleq CA^j B, \quad j \geq 0 \quad (24)$$

$$H_{yw,-1} \triangleq D_2, \quad H_{yw,j} \triangleq CA^j D_1, \quad j \geq 0 \quad (25)$$

$$H_{zu,-1} \triangleq E_2, \quad H_{zu,j} \triangleq E_1 A^j B, \quad j \geq 0 \quad (26)$$

$$H_{zw,-1} \triangleq 0, \quad H_{zw,j} \triangleq E_1 A^j D_1, \quad j \geq 0 \quad (27)$$

the ARMA model of (1)–(3) is given by

$$\begin{aligned} z(k) = & \sum_{j=1}^n -\alpha_j z(k - \mu - j + 1) \\ & + \sum_{j=1}^{\mu} H_{zw,j-2} w(k - j + 1) \\ & + \sum_{j=1}^n B_{zw,j} w(k - \mu - j + 1) \\ & + \sum_{j=1}^{\mu} H_{zu,j-2} u(k - j + 1) \\ & + \sum_{j=1}^n B_{zu,j} u(k - \mu - j + 1) \end{aligned} \quad (28)$$

$$\begin{aligned} y(k) = & \sum_{j=1}^n -\alpha_j y(k - \mu - j + 1) \\ & + \sum_{j=1}^{\mu} H_{yw,j-2} w(k - j + 1) \\ & + \sum_{j=1}^n B_{yw,j} w(k - \mu - j + 1) \\ & + \sum_{j=1}^{\mu} H_{yu,j-2} u(k - j + 1) \\ & + \sum_{j=1}^n B_{yu,j} u(k - \mu - j + 1) \end{aligned} \quad (29)$$

where $\alpha_j \in \mathcal{R}$, $B_{zw,j}, H_{zw,j} \in \mathcal{R}^{l_z \times m_w}$, $B_{zu,j}, H_{zu,j} \in \mathcal{R}^{l_z \times m_u}$, $B_{yw,j}, H_{yw,j} \in \mathcal{R}^{l_y \times m_w}$, and $B_{yu,j}, H_{yu,j} \in \mathcal{R}^{l_y \times m_u}$.

Next, define the *extended performance vector* $Z(k)$, the *extended measurement vector* $Y(k)$ and the *extended control vector* $U(k)$ by

$$\begin{aligned} Z(k) & \triangleq \begin{bmatrix} z(k) \\ \vdots \\ z(k-p+1) \end{bmatrix}, \quad Y(k) \triangleq \begin{bmatrix} y(k) \\ \vdots \\ y(k-p+1) \end{bmatrix}, \\ U(k) & \triangleq \begin{bmatrix} u(k) \\ \vdots \\ u(k-p_c+1) \end{bmatrix} \end{aligned} \quad (30)$$

where $p_c \triangleq \mu + n + p - 1$, and the *ARMA regressor vectors* $\Phi_{zw}(k)$ and $\Phi_{yw}(k)$ by

$$\begin{aligned} \Phi_{zw}(k) & \triangleq \begin{bmatrix} z(k-\mu) \\ \vdots \\ z(k-\mu-p-n+2) \\ w(k) \\ \vdots \\ w(k-\mu-p-n+2) \end{bmatrix} \\ \Phi_{yw}(k) & \triangleq \begin{bmatrix} y(k-\mu) \\ \vdots \\ y(k-\mu-p-n+2) \\ w(k) \\ \vdots \\ w(k-\mu-p-n+2) \end{bmatrix}. \end{aligned} \quad (31)$$

Furthermore, define the block-Toeplitz *ARMA weight matrices* $W_{zw} \in \mathcal{R}^{p_l z \times [(n+p-1)l_z + (\mu+n+p-1)m_w]}$ and $W_{yw} \in \mathcal{R}^{p_l y \times [(n+p-1)l_y + (\mu+n+p-1)m_w]}$ by (32)–(35) shown at the bottom of the next page and the block-Toeplitz *ARMA control matrices* $B_{zu} \in \mathcal{R}^{p_l z \times p_c m_u}$ and $B_{yu} \in \mathcal{R}^{p_l y \times p_c m_u}$ by

Then (28) and (29) can be written in the form

$$Z(k) = W_{zw} \Phi_{zw}(k) + B_{zu} U(k) \quad (36)$$

$$Y(k) = W_{yw} \Phi_{yw}(k) + B_{yu} U(k) \quad (37)$$

which is the ARMA/Toeplitz model of (1)–(3).

V. ADAPTIVE DISTURBANCE REJECTION ALGORITHM

In this section we formulate an adaptive disturbance rejection algorithm for the TITO system represented by (36) and (37). We use a strictly proper controller in ARMA form of order n_c with μ_c Markov parameters, so that, analogous to (19), the control $u(k)$ is given by

$$\begin{aligned} u(k) = & \sum_{j=1}^{n_c} -\alpha_{c,j}(k) u(k - \mu_c - j + 1) \\ & + \sum_{j=1}^{\mu_c-1} H_{c,j-1}(k) y(k - j + 1) \\ & + \sum_{j=1}^{n_c} B_{c,j}(k) y(k - \mu_c - j + 1) \end{aligned} \quad (38)$$

where $H_{c,j} \in \mathcal{R}^{m_u \times l_y}$, $j = 1, 2, \dots$, are the Markov parameters of the controller. Next, define the *controller parameter block vector* $\theta(k) \in \mathcal{R}^{m_u \times [n_c m_u + (n_c + \mu_c - 1)l_y]}$ by

$$\theta(k) \triangleq [-\alpha_{c,1}(k)I_{m_u} \cdots -\alpha_{c,n_c}(k)I_{m_u} \ H_{c,0}(k) \cdots H_{c,\mu_c-2}(k) \ \mathcal{B}_{c,1}(k) \cdots \mathcal{B}_{c,n_c}(k)]. \quad (39)$$

Now from (30) and (38) it follows that $u(k)$ and $U(k)$ are given by

$$u(k) = \theta(k)R_1\Phi_{uy}(k) \quad (40)$$

and

$$U(k) = \sum_{i=1}^{p_c} L_i \theta(k-i+1)R_i \Phi_{uy}(k) \quad (41)$$

where

$$\Phi_{uy}(k) \triangleq \begin{bmatrix} u(k-\mu_c) \\ \vdots \\ u(k-\mu_c-n_c-p_c+2) \\ y(k-1) \\ \vdots \\ y(k-\mu_c-n_c-p_c+2) \end{bmatrix} \quad (42)$$

and where

$$L_i \triangleq \begin{bmatrix} 0_{(i-1)m_u \times m_u} \\ I_{m_u} \\ 0_{(p_c-i)m_u \times m_u} \end{bmatrix} \in \mathcal{R}^{p_c m_u \times m_u} \quad (43)$$

where R_i is defined in (44) at the bottom of the next page with $q_1 \triangleq n_c m_u$ and $q_2 \triangleq (n_c + \mu_c - 1)l_y$. Thus, from (36) and (41) we obtain

$$Z(k) = W_{zw}\Phi_{zw}(k) + B_{zu} \sum_{i=1}^{p_c} L_i \theta(k-i+1)R_i \Phi_{uy}(k). \quad (45)$$

$$W_{zw} \triangleq \begin{bmatrix} -\alpha_1 I_z & \cdots & -\alpha_n I_z & 0_{l_z} & \cdots & 0_{l_z} & H_{zw,-1} & \cdots & H_{zw,\mu-2} & \mathcal{B}_{zw,1} & \cdots & \mathcal{B}_{zw,n} & 0_{l_z \times m_w} & \cdots & 0_{l_z \times m_w} \\ 0_{l_z} & \ddots & & \ddots & \ddots & \vdots & 0_{l_z \times m_w} & \ddots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0_{l_z} & \vdots & \ddots & \ddots & & \ddots & \ddots & \ddots & \ddots & 0_{l_z \times m_w} \\ 0_{l_z} & \cdots & 0_{l_z} & -\alpha_1 I_z & \cdots & -\alpha_n I_z & 0_{l_z \times m_w} & \cdots & 0_{l_z \times m_w} & H_{zw,-1} & \cdots & H_{zw,\mu-2} & \mathcal{B}_{zw,1} & \cdots & \mathcal{B}_{zw,n} \end{bmatrix} \quad (32)$$

$$W_{yw} \triangleq \begin{bmatrix} -\alpha_1 I_y & \cdots & -\alpha_n I_y & 0_{l_y} & \cdots & 0_{l_y} & H_{yw,-1} & \cdots & H_{yw,\mu-2} & \mathcal{B}_{yw,1} & \cdots & \mathcal{B}_{yw,n} & 0_{l_y \times m_w} & \cdots & 0_{l_y \times m_w} \\ 0_{l_y} & \ddots & & \ddots & \ddots & \vdots & 0_{l_y \times m_w} & \ddots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0_{l_y} & \vdots & \ddots & \ddots & & \ddots & \ddots & \ddots & \ddots & 0_{l_y \times m_w} \\ 0_{l_y} & \cdots & 0_{l_y} & -\alpha_1 I_y & \cdots & -\alpha_n I_y & 0_{l_y \times m_w} & \cdots & 0_{l_y \times m_w} & H_{yw,-1} & \cdots & H_{yw,\mu-2} & \mathcal{B}_{yw,1} & \cdots & \mathcal{B}_{yw,n} \end{bmatrix} \quad (33)$$

$$B_{zu} \triangleq \begin{bmatrix} H_{zu,-1} & \cdots & H_{zu,\mu-2} & \mathcal{B}_{zu,1} & \cdots & \mathcal{B}_{zu,1} & 0_{l_z \times m_u} & \cdots & 0_{l_z \times m_u} \\ 0_{l_z \times m_u} & \ddots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \ddots & \ddots & 0_{l_z \times m_u} \\ 0_{l_z \times m_u} & \cdots & 0_{l_z \times m_u} & H_{zu,-1} & \cdots & H_{zu,\mu-2} & \mathcal{B}_{zu,1} & \cdots & \mathcal{B}_{zu,n} \end{bmatrix} \quad (34)$$

$$B_{yu} \triangleq \begin{bmatrix} H_{yu,-1} & \cdots & H_{yu,\mu-2} & \mathcal{B}_{yu,1} & \cdots & \mathcal{B}_{yu,n} & 0_{l_y \times m_u} & \cdots & 0_{l_y \times m_u} \\ 0_{l_y \times m_u} & \ddots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \ddots & \ddots & 0_{l_y \times m_u} \\ 0_{l_y \times m_u} & \cdots & 0_{l_y \times m_u} & H_{yu,-1} & \cdots & H_{yu,\mu-2} & \mathcal{B}_{yu,1} & \cdots & \mathcal{B}_{yu,n} \end{bmatrix}. \quad (35)$$

Next, we derive an update law for the controller parameter block vector $\theta(k)$. To do this, we consider a retrospective performance cost function that evaluates the performance of the current value of $\theta(k)$ based upon the behavior of the system during the previous p_c steps. Therefore, we define the *retrospective performance* $\hat{Z}(k)$ by

$$\hat{Z}(k) \triangleq W_{zw} \Phi_{zw}(k) + B_{zu} \sum_{i=1}^{p_c} L_i \theta(k) R_i \Phi_{uy}(k) \quad (46)$$

which has the same form as (45) but with $\theta(k-i+1)$ replaced by the current controller parameter block vector $\theta(k)$. Using (46) we define the *retrospective performance cost function*

$$J(k) = \frac{1}{2} \hat{Z}^T(k) \hat{Z}(k). \quad (47)$$

Lemma 1: The gradient of $J(k)$ with respect to $\theta(k)$ is given by

$$\frac{\partial J(k)}{\partial \theta(k)} = \sum_{i=1}^{p_c} L_i^T B_{zu}^T \hat{Z}(k) \Phi_{uy}^T(k) R_i^T. \quad (48)$$

Proof: See Appendix A. \square

Since $w(k)$ is not available, which implies that $\Phi_{zw}(k)$ is unknown, $\hat{Z}(k)$ cannot be calculated from (46). However, it follows from (36) and (46) that

$$\hat{Z}(k) = Z(k) - B_{zu} \left(U(k) - \sum_{i=1}^{p_c} L_i \theta(k) R_i \Phi_{uy}(k) \right) \quad (49)$$

which can be used to evaluate (48).

The gradient (48) is used in the update law

$$\theta(k+1) = \theta(k) - \eta(k) \frac{\partial J(k)}{\partial \theta(k)} \quad (50)$$

where $\eta(k)$ is the *adaptive step size*. To determine the adaptive step size $\eta(k)$, we make the following assumption which is analogous to the assumption given in [10, pp. 281–282].

Assumption 1: There exists $\theta^* \in \mathcal{R}^{m_u \times [n_c m_u + (n_c + \mu_c - 1) l_y]}$ that minimizes $J(k)$ for all k . Under Assumption 1, we define the *desired performance*

$$\hat{Z}^*(k) \triangleq W_{zw} \Phi_{zw}(k) + B_{zu} \sum_{i=1}^{p_c} L_i \theta^* R_i \Phi_{uy}(k) \quad (51)$$

the *error matrix*

$$E(k) \triangleq \theta^* - \theta(k) \quad (52)$$

the *performance error*

$$\varepsilon(k) \triangleq \hat{Z}^*(k) - \hat{Z}(k) \quad (53)$$

and the *error matrix cost function*

$$\mathcal{J}(k, \eta(k)) \triangleq \|E(k+1)\|_F^2 - \|E(k)\|_F^2. \quad (54)$$

Our goal is to determine $\eta(k)$ such that $\|E(k)\|_F^2$ is decreasing, that is, $\mathcal{J}(k, \eta(k))$ is negative. For convenience in stating the following result, we define the *optimal adaptive step size*

$$\eta_{\text{opt}}(k) \triangleq \frac{\|\varepsilon(k)\|_2^2}{\left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F^2}. \quad (55)$$

Theorem 1: Consider the update law (50) and suppose Assumption 1 is satisfied. Furthermore, let $k \geq 0$ and assume that $(\partial J(k)/\partial \theta(k)) \neq 0$. Then

$$\mathcal{J}(k, \eta(k)) < 0 \quad (56)$$

if and only if

$$0 < \eta(k) < 2\eta_{\text{opt}}(k). \quad (57)$$

In particular, $\eta(k) = \eta_{\text{opt}}(k)$ minimizes $\mathcal{J}(k, \eta(k))$. Finally,

$$\mathcal{J}(k, \eta_{\text{opt}}(k)) = - \frac{\|\varepsilon(k)\|_2^4}{\left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F^2}. \quad (58)$$

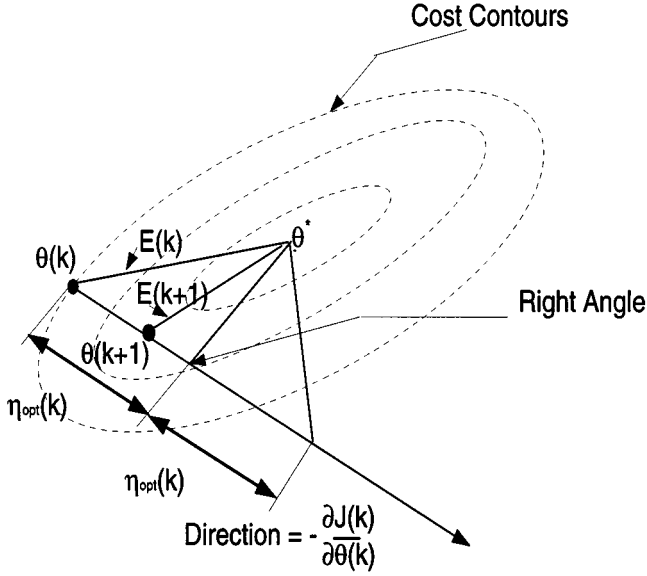
Proof: See Appendix B. \square

A geometrical interpretation of Theorem 1 is now presented. Using Fig. 3 for reference, the objective of the algorithm is to move the controller parameter block vector $\theta(k)$ closer to the optimal controller parameter block vector θ^* . The direction to move is the negative of the gradient $\partial J(k)/\partial \theta(k)$ given by (48) which is obtained from the retrospective performance cost function (47). The distance to move is determined by the adaptive step size $\eta(k)$. Theorem 1 states that the step size $\eta_{\text{opt}}(k)$ moves $\theta(k)$ to the point closest to θ^* along the negative gradient direction and, at this point, the vectors $E(k+1)$ and $-(\partial J(k)/\partial \theta(k))$ are at right angles.

In practice $\eta_{\text{opt}}(k)$ given by (55) is not computable since $\varepsilon(k)$ is not available. Hence, we define the *implementable adaptive step sizes* $\eta_1(k)$, $\eta_2(k)$, and $\eta_3(k)$ by

$$\eta_1(k) \triangleq \frac{1}{\left[\sum_{i=1}^{p_c} \|R_i \Phi_{uy}(k)\|_2 \bar{\sigma}(B_{zu} L_i) \right]^2} \quad (59)$$

$$R_i \triangleq \begin{bmatrix} 0_{q_1 \times (i-1)m_u} & I_{q_1 \times q_1} & 0_{q_1 \times (p_c-i)m_u} & 0_{q_1 \times (i-1)l_y} & 0_{q_1 \times q_2} & 0_{q_1 \times (p_c-i)l_y} \\ 0_{q_2 \times (i-1)m_u} & 0_{q_2 \times q_1} & 0_{q_2 \times (p_c-i)m_u} & 0_{q_2 \times (i-1)l_y} & I_{q_2 \times q_2} & 0_{q_2 \times (p_c-i)l_y} \end{bmatrix} \in \mathcal{R}^{[n_c m_u + (n_c + \mu_c - 1) l_y] \times [(n_c + p_c - 1) m_u + (n_c + \mu_c + p_c - 2) l_y]} \quad (44)$$


 Fig. 3. Cost contours for $J(k)$.

$$\eta_2(k) \triangleq \frac{1}{\|\Phi_{uy}(k)\|_2^2 \left[\sum_{i=1}^{p_c} \bar{\sigma}(B_{zu}L_i) \right]^2} \quad (60)$$

$$\eta_3(k) \triangleq \frac{1}{p_c \|B_{zu}\|_F^2 \|\Phi_{uy}(k)\|_2^2} \quad (61)$$

where $\bar{\sigma}(B_{zu}L_i)$ denotes the maximum singular value of the matrix $B_{zu}L_i$. Note that if B_{zu} is known, then $\eta_1(k)$, $\eta_2(k)$ or $\eta_3(k)$ can be calculated and used to implement (50). The following result shows that $\eta_1(k)$, $\eta_2(k)$, and $\eta_3(k)$ satisfy the requirements of Theorem 1. We note that although $\eta_2(k) \leq \eta_1(k)$, the computation of $\eta_2(k)$ is less burdensome.

Proposition 1: The implementable step sizes $\eta_1(k)$, $\eta_2(k)$, and $\eta_3(k)$ satisfy

$$0 < \eta_3(k) \leq \eta_2(k) \leq \eta_1(k) \leq \eta_{opt}(k). \quad (62)$$

Proof: See Appendix C. \square

Finally, we show that the update law (50) with step sizes $\eta_1(k)$, $\eta_2(k)$, and $\eta_3(k)$ drives $\hat{Z}(k)$ to $\hat{Z}^*(k)$ as k tends to infinity if $\{\Phi_{uy}(k)\}_{k=0}^{\infty}$ is bounded.

Proposition 2: Suppose Assumption 1 is satisfied, let $(\partial J(k)/\partial \theta(k)) \neq 0$ for all $k \geq 0$ and let $\eta(k) = \eta_1(k)$, $\eta_2(k)$ or $\eta_3(k)$. If $\{\Phi_{uy}(k)\}_{k=0}^{\infty}$ is bounded, then

$$\lim_{k \rightarrow \infty} \varepsilon(k) = 0. \quad (63)$$

Proof: See Appendix D. \square

VI. CONTROLLER IMPLEMENTATION

The steps involved in implementing the adaptive algorithm are as follows.

- 1) Obtain the matrix B_{zu} by using the identification algorithm of [20], [21] or by calculation from an ARMA or a state-space representation of G_{zu} .
- 2) Calculate the control signal $u(k)$ from the controller parameter matrix $\theta(k)$ and the vector $\Phi_{uy}(k)$ using (40).
- 3) Using the signals $u(k)$, $z(k)$ and $y(k)$ update the vectors $\hat{Z}(k)$ and $\Phi_{uy}(k)$ as defined in (42) and (49).
- 4) Calculate the gradient $\partial J(k)/\partial \theta(k)$ using (48).
- 5) Calculate the implementable adaptive step size $\eta_1(k)$, $\eta_2(k)$ or $\eta_3(k)$ from (59)–(61).
- 6) Update the controller parameter matrix $\theta(k)$ using (50).

Steps 1)–5) are performed at each time step k .

We observe that of the four transfer matrices G_{zw} , G_{zu} , G_{yw} , and G_{yu} in the multiple input–multiple output (MIMO) standard problem, the algorithm described above requires that we identify only the numerator of one transfer matrix, namely, G_{zu} . The signals that we require to be measured are $y(k)$ and $z(k)$.

VII. EXPERIMENTAL RESULTS

Experimental demonstration of the ARMARKOV adaptive disturbance algorithm is performed on an acoustic duct of circular cross section. The duct is 80 in long and has a diameter of 4 in. The disturbance speaker (w) is located at one end of the duct and the measurement microphone (y) is located 4 in from the same end of the duct. The performance microphone (z) is positioned 6 in from the other end of the duct while the control speaker (u) is placed 16 in away from that end of the duct. The signals from the two microphones are amplified by a dbx 760x microphone preamplifier while the control signal is amplified by an Alesis RA-100 amplifier. Both speakers are Radio Shack 6-in woofers. The experimental setup is shown in Fig. 4.

The algorithm is tested on four types of disturbances, namely a single-tone disturbance (139.65 Hz), a two-tone disturbance (135.74 Hz and 160.4 Hz), band-limited white noise (up to 390 Hz) and AM radio noise. The algorithm uses $n = 4$ and $\mu = 12$ for the matrix B_{zu} , and $n_c = 2$, $\mu_c = 10$, $p = 2$, and the step-size $\eta_3(k)$ for the adaptive controller. The controller is implemented on a dSPACE ds1102 real-time control board running a TMS320C30 DSP processor at a sampling frequency of 800 Hz. The microphone signals are processed through a four-pole Ithaco low-pass filter model DL 4302 that rolls off at 315 Hz. The tonal and band-limited white noise disturbances are generated by a Stanford Research Systems 770 FFT network analyzer and amplified by an Optimus STA-825 stereo receiver.

Fig. 5 shows the open-loop and closed-loop frequency domain performance with a single-tone disturbance. Disturbance attenuation of over 40 dB is achieved with convergence in approximately 1 s. Although the disturbance signal is a pure tone, speaker nonlinearities produce harmonics which appear on the frequency response plot along with ambient and measurement noise. The algorithm provides the same level of attenuation by adaptation when the frequency of the disturbance tone is changed or swept while the algorithm is running.

Fig. 6 shows the open-loop and closed-loop performance with a two-tone disturbance. In this case, disturbance attenuation over 35 dB is observed. Fig. 7 shows the open-loop

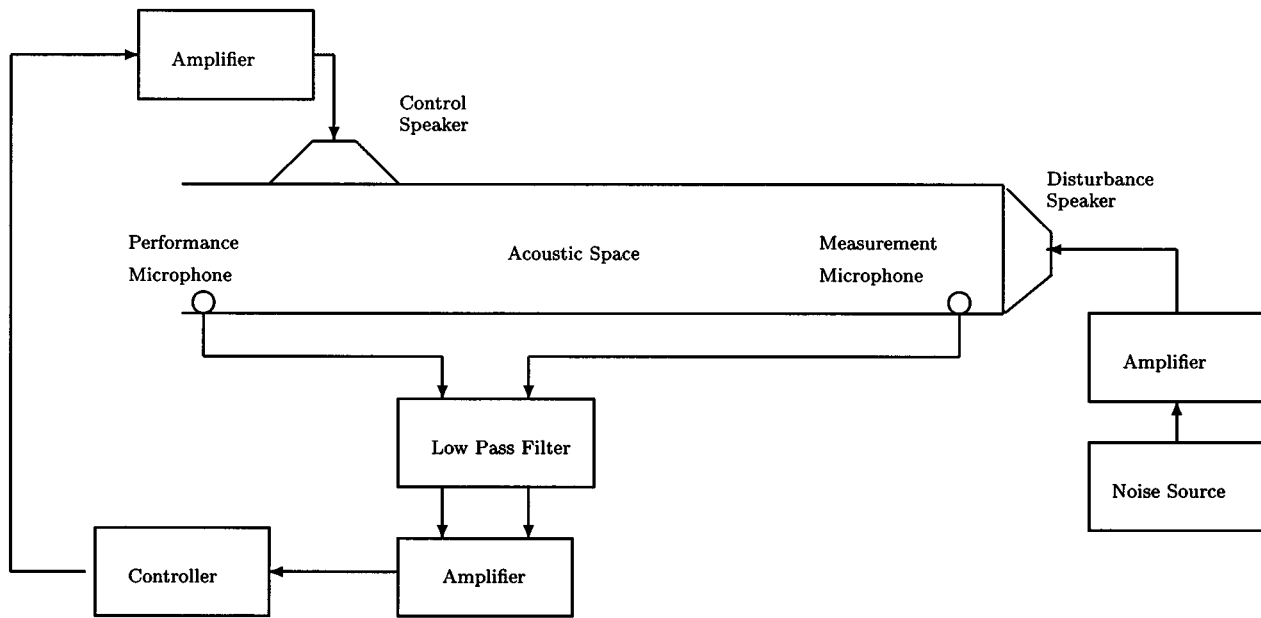


Fig. 4. Experimental setup for noise suppression in an acoustic duct.

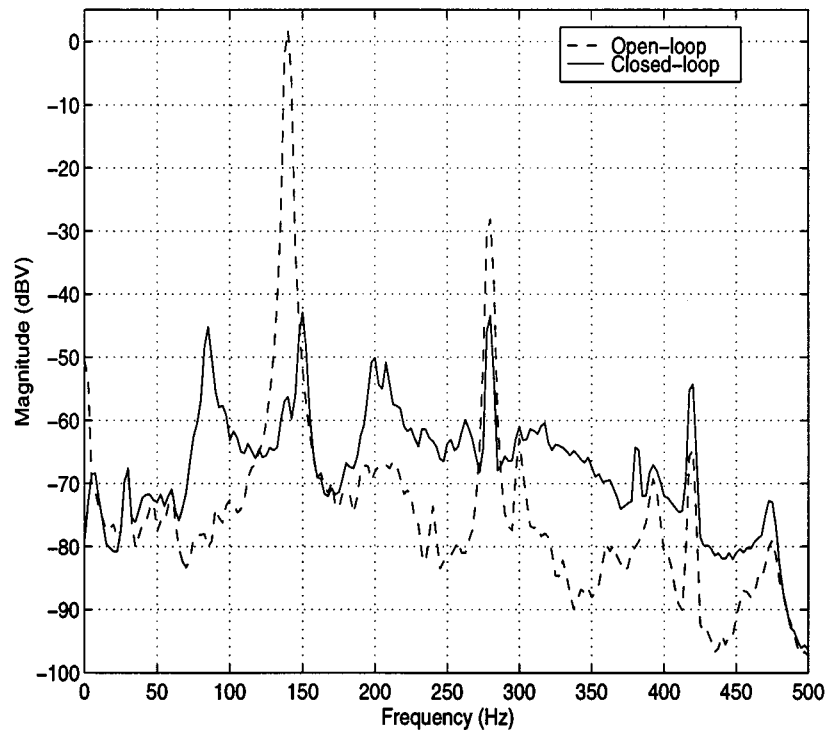


Fig. 5. Open-loop and closed-loop frequency domain performance with a single-tone disturbance at 139.65 Hz.

and closed-loop magnitude plots of the transfer function from disturbance to performance with a white noise disturbance, and noise suppression of up to 15 dB is observed over a frequency range from 0 to 300 Hz. Finally, Fig. 8 shows the open-loop and closed-loop response with an AM radio disturbance. Noise reduction levels of up to 40 dB are observed over the frequency range 0 to 300 Hz.

The performance of the algorithm was also experimentally compared to the filtered-x FIR LMS algorithm (FXLMS) and

the filtered-u IIR LMS (FURLMS) [2]. A brief description of these algorithms is given in [26]. Each algorithm was tested on the testbed described above with the following disturbances: a single tone at 115 Hz, dual tones at 115 Hz and 125 Hz, band-limited white noise, and recorded fan noise. The parameters used for the algorithms are in Table I. The step sizes used for the FXLMS and FURLMS algorithms were determined by trial and error to obtain the fastest convergence rate with consistent stability. The ARMARKOV/Toeplitz algorithm was

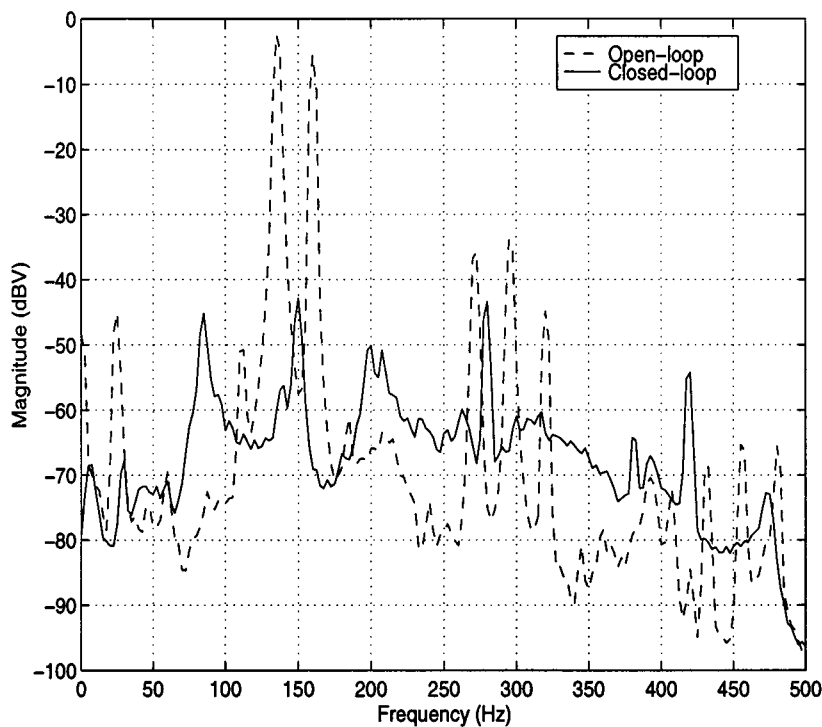


Fig. 6. Open-loop and closed-loop frequency domain performance with a two-tone disturbance at 135.74 and 160.4 Hz.

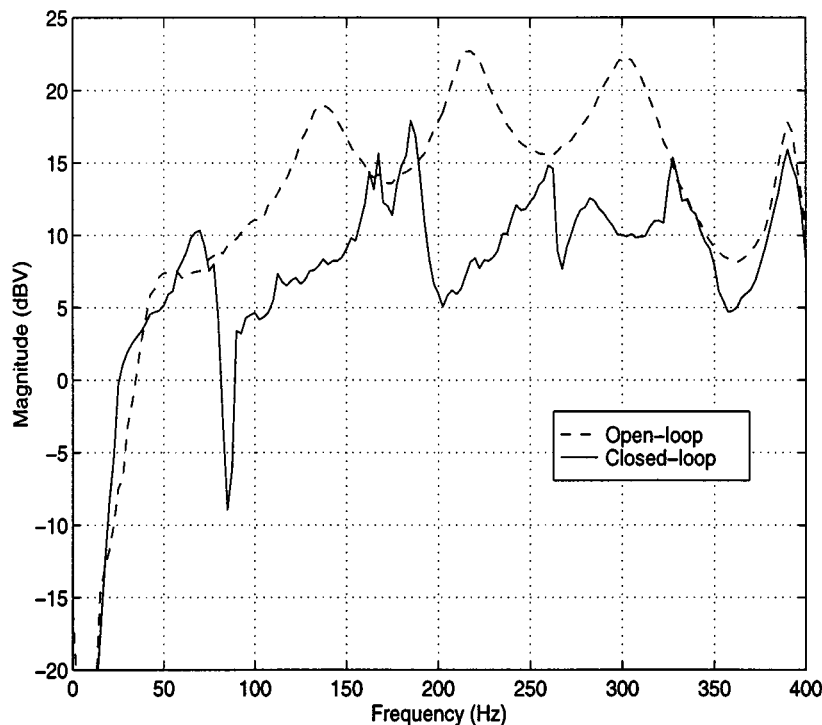


Fig. 7. Open-loop and closed-loop performance with band-limited white noise.

implemented using $\eta_3(k)$ as the step size. All of the algorithms were successful in rejecting computer-generated single and dual tones to varying degrees. Only the ARMARKOV/Toeplitz algorithm was capable of attenuating the computer-generated band-limited white noise and the fan disturbance. Also, the ARMARKOV/Toeplitz algorithm converged faster in general as a result of the use of an adaptive step size. The results, sum-

marized in Table II, indicate that the ARMARKOV/Toeplitz algorithm has better stability and performance characteristics.

VIII. DISCUSSION

In this paper, we developed an adaptive disturbance rejection algorithm based on ARMARKOV/Toeplitz models. The al-

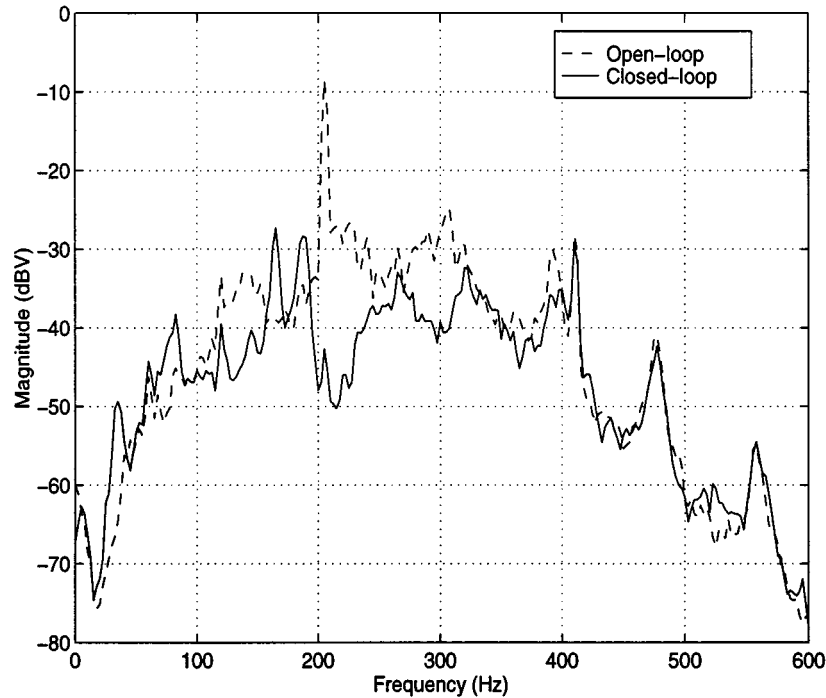


Fig. 8. Open-loop and closed-loop performance with AM radio disturbance.

TABLE I
ALGORITHM PARAMETERS

Algorithm	Free Parameters	Step Size	Type
FXLMS	30	$\mu = 10^{-5}$ (fixed)	FIR
FURLMS	58	$\mu = 10^{-5}$ (fixed)	IIR
ARM/Toep	14	automatic	IIR

TABLE II
DISTURBANCE ATTENUATION COMPARISON

Algorithm	Single Tone	Dual Tones	White Noise	Fan Noise
FXLMS	29.3dB	26.5 dB	unstable	unstable
FURLMS	56.3 dB	24.3 dB	unstable	unstable
ARM/Toep	89.5 dB	24.9 dB	6.5 dB (avg)	60.2 dB

gorithm is gradient based with an adaptive step size. In contrast to LMS algorithms, the ARMARKOV/Toeplitz algorithm has three distinctive features, namely, the use of ARMARKOV system representations (36), (37); a retrospective performance cost function obtained by using the current controller over a past window of data (47); and an implementable adaptive step size (61). The retrospective cost function allows the calculation of an exact gradient for the system without resorting to prefiltering of $y(k)$ by G_{zu} as in filtered-x methods. This prefiltering, which is necessary to account for G_{zu} in the gradient, limits the adaptation rate and thus adversely affects the convergence of LMS methods. The use of the implementable adaptive step size guarantees that the updated controller $\theta(k+1)$ moves closer to the desired controller θ^* as stated in Theorem 1, thus removing the need to choose a step size based on trial and error as is done for LMS algorithms [2], [5], [7]. The algorithm was applied to the

problem of active noise control in an acoustic duct, and experimental results showed that the algorithm is effective in rejecting both narrow-band and broad-band disturbances with minimal plant modeling.

APPENDIX A

Proof of Lemma 1: From (46) and (47) it follows that

$$J(k) = \frac{1}{2} [W_{zw}\Phi_{zw}(k) + B_{zu} \sum_{i=1}^{p_c} L_i \theta(k) R_i \Phi_{uy}(k)]^T \cdot [W_{zw}\Phi_{zw}(k) + B_{zu} \sum_{i=1}^{p_c} L_i \theta(k) R_i \Phi_{uy}(k)]. \quad (64)$$

Using matrix derivative formulas, it follows from (64) and (46) that

$$\begin{aligned} \frac{\partial J(k)}{\partial \theta(k)} &= \sum_{i=1}^{p_c} L_i^T B_{zu}^T W_{zw} \Phi_{zw}(k) \Phi_{uy}^T(k) R_i^T \\ &+ \sum_{i=1}^{p_c} L_i^T B_{zu}^T B_{zu} \left(\sum_{j=1}^{p_c} L_j \theta(k) R_j \right) \cdot \Phi_{uy}(k) \Phi_{uy}^T(k) R_i^T \\ &= \sum_{i=1}^{p_c} L_i^T B_{zu}^T \left[W_{zw} \Phi_{zw}(k) + B_{zu} \sum_{j=1}^{p_c} L_j \theta(k) R_j \Phi_{uy}(k) \right] \Phi_{uy}^T(k) R_i^T \\ &= \sum_{i=1}^{p_c} L_i^T B_{zu}^T \hat{Z}(k) \Phi_{uy}^T(k) R_i^T. \end{aligned} \quad (65)$$

□

APPENDIX B

Proof of Theorem 1: From (46), (51), and (53) it follows that

$$\varepsilon(k) = B_{zu} \sum_{i=1}^{p_c} L_i [\theta^* - \theta(k)] R_i \Phi_{uy}(k). \quad (66)$$

Using (52), (66) can be written as

$$\varepsilon(k) = B_{zu} \sum_{i=1}^{p_c} L_i E(k) R_i \Phi_{uy}(k). \quad (67)$$

By Assumption 1, θ^* minimizes $J(k)$, and thus it follows from (48) that

$$\left. \frac{\partial J(k)}{\partial \theta(k)} \right|_{\theta(k)=\theta^*} = \sum_{i=1}^{p_c} L_i^T B_{zu}^T \hat{Z}^*(k) \Phi_{uy}^T(k) R_i^T = 0. \quad (68)$$

Subtracting (68) from (48) and substituting $\varepsilon(k)$ from (53) into the resulting equation yields

$$\begin{aligned} \frac{\partial J(k)}{\partial \theta(k)} &= \sum_{i=1}^{p_c} L_i^T [B_{zu}^T (\hat{Z}(k) - \hat{Z}^*(k)) \Phi_{uy}^T(k)] R_i^T \\ &= - \sum_{i=1}^{p_c} L_i^T [B_{zu}^T \varepsilon(k) \Phi_{uy}^T(k)] R_i^T. \end{aligned} \quad (69)$$

Next, using (50) and (52) we obtain

$$E(k+1) = E(k) + \eta(k) \frac{\partial J(k)}{\partial \theta(k)} \quad (70)$$

and thus

$$\begin{aligned} \|E(k+1)\|_F^2 - \|E(k)\|_F^2 &= 2\eta(k) \text{tr} \left(E(k) \frac{\partial J(k)}{\partial \theta(k)}^T \right) \\ &\quad + \eta^2(k) \left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F^2. \end{aligned} \quad (71)$$

Using (67) and (69) we obtain

$$\begin{aligned} &\text{tr} \left(E(k) \frac{\partial J(k)}{\partial \theta(k)}^T \right) \\ &= -\text{tr} \left(E(k) \sum_{i=1}^{p_c} R_i \Phi_{uy}(k) \varepsilon^T(k) B_{zu} L_i \right) \\ &= - \sum_{i=1}^{p_c} (\varepsilon^T(k) B_{zu} L_i E(k) R_i \Phi_{uy}(k)) \\ &= -\varepsilon^T(k) \left(B_{zu} \sum_{i=1}^{p_c} L_i E(k) R_i \Phi_{uy}(k) \right) \\ &= -\|\varepsilon(k)\|_2^2. \end{aligned} \quad (72)$$

Thus, (54), (71), and (72) imply that

$$\mathcal{J}(k, \eta(k)) = -2\eta(k) \|\varepsilon(k)\|_2^2 + \eta^2(k) \left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F^2. \quad (73)$$

Now, from (73) it follows that $\mathcal{J}(k, \eta(k)) < 0$ if and only if

$$\eta(k) < \frac{2\|\varepsilon(k)\|_2^2}{\left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F^2} = 2\eta_{\text{opt}}(k) \quad (74)$$

which proves the first statement of the theorem.

To prove the second statement of the theorem we note from (73) that

$$\begin{aligned} \mathcal{J}(k, \eta(k)) &= \eta^2(k) \left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F^2 - 2\|\varepsilon(k)\|_2^2 \eta(k) \end{aligned} \quad (75)$$

$$= [\eta^2(k) - 2\eta(k)\eta_{\text{opt}}(k)] \left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F^2 \quad (76)$$

$$= [(\eta(k) - \eta_{\text{opt}}(k))^2 - \eta_{\text{opt}}^2(k)] \left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F^2. \quad (77)$$

Since the quadratic function $(\eta(k) - \eta_{\text{opt}}(k))^2 - \eta_{\text{opt}}^2(k)$ achieves its minimum at $\eta(k) = \eta_{\text{opt}}(k)$, it follows from (77) that $\mathcal{J}(p, \eta(k))$ is minimized by $\eta(k) = \eta_{\text{opt}}(k)$. Substituting (55) into (73) yields (58). \square

APPENDIX C

Proof of Proposition 1: First, from (69) we note that

$$\left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F = \left\| - \sum_{i=1}^{p_c} R_i \Phi_{uy}(k) \varepsilon(k)^T B_{zu} L_i \right\|_F \quad (78)$$

$$\leq \sum_{i=1}^{p_c} \|R_i \Phi_{uy}(k) \varepsilon(k)^T B_{zu} L_i\|_F \quad (79)$$

$$\leq \sum_{i=1}^{p_c} \|R_i \Phi_{uy}(k)\|_2 \|\varepsilon(k)\|_2 \bar{\sigma}(B_{zu} L_i) \quad (80)$$

$$= \|\varepsilon(k)\|_2 \sum_{i=1}^{p_c} \|R_i \Phi_{uy}(k)\|_2 \bar{\sigma}(B_{zu} L_i) \quad (81)$$

$$\leq \|\varepsilon(k)\|_2 \|\Phi_{uy}(k)\|_2 \sum_{i=1}^{p_c} \bar{\sigma}(B_{zu} L_i) \quad (82)$$

$$\leq \|\varepsilon(k)\|_2 \|\Phi_{uy}(k)\|_2 \sum_{i=1}^{p_c} \|B_{zu}\|_F. \quad (83)$$

Thus, from (81)–(83) it follows that

$$\frac{1}{\left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F^2} \geq \frac{1}{\|\varepsilon(k)\|_2^2 \left[\sum_{i=1}^{p_c} \|R_i \Phi_{uy}(k)\|_2 \bar{\sigma}(B_{zu} L_i) \right]^2} \quad (84)$$

$$\geq \frac{1}{\|\varepsilon(k)\|_2^2 \|\Phi_{uy}(k)\|_2^2 \left[\sum_{i=1}^{p_c} \bar{\sigma}(B_{zu} L_i) \right]^2} \quad (85)$$

$$\geq \frac{1}{p_c \|\varepsilon(k)\|_2^2 \|\Phi_{uy}(k)\|_2^2 \|B_{zu}\|_F^2} \quad (86)$$

and hence

$$\frac{\|\varepsilon(k)\|_2^2}{\left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F^2} \geq \frac{1}{\left[\sum_{i=1}^{p_c} \|R_i \Phi_{uy}(k)\|_2 \bar{\sigma}(B_{zu} L_i) \right]^2} \quad (87)$$

$$\geq \frac{1}{\|\Phi_{uy}(k)\|_2^2 \left[\sum_{i=1}^{p_c} \bar{\sigma}(B_{zu} L_i) \right]^2} \quad (88)$$

$$\geq \frac{1}{p_c \|\varepsilon(k)\|_2^2 \|\Phi_{uy}(k)\|_2^2 \|B_{zu}\|_F^2}. \quad (89)$$

Using (55), (59)–(61), and (87)–(89) it follows that $\eta_1(k)$, $\eta_2(k)$ and $\eta_3(k)$ satisfy (62). \square

APPENDIX D

Proof of Proposition 2: Let $j = 1, 2, 3$ and define

$$\Delta(k) \triangleq \frac{1}{\sqrt{\eta_j(k)}}. \quad (90)$$

Then, from (84)–(86) it follows that

$$\left\| \frac{\partial J(k)}{\partial \theta(k)} \right\|_F \leq \|\varepsilon(k)\|_2 \Delta(k). \quad (91)$$

Now (59)–(61), (73), and (91) imply that

$$\mathcal{J}(k, \eta_j(k)) \leq -2 \frac{\|\varepsilon(k)\|_2^2}{\Delta(k)^2} + \frac{\|\varepsilon(k)\|_2^2 \Delta(k)^2}{\Delta(k)^4} \quad (92)$$

$$\leq -\frac{\|\varepsilon(k)\|_2^2}{\Delta(k)^2}. \quad (93)$$

Using (54) and (93) we obtain

$$\|E(k+1)\|_F^2 - \|E(k)\|_F^2 \leq -\frac{\|\varepsilon(k)\|_2^2}{\Delta(k)^2}. \quad (94)$$

Next

$$\begin{aligned} \|E(0)\|_F^2 &> \|E(0)\|_F^2 - \|E(r+1)\|_F^2 \\ &= \sum_{k=0}^r (\|E(k)\|_F^2 - \|E(k+1)\|_F^2). \end{aligned} \quad (95)$$

Substituting (94) into (95) yields

$$\sum_{k=0}^r \frac{\|\varepsilon(k)\|_2^2}{\Delta(k)^2} < \|E(0)\|_F^2. \quad (96)$$

Since $\{\Phi_{uy}(k)\}_{k=0}^{\infty}$ is assumed to be bounded, there exists $\beta > 0$ such that $\Delta(k)^2 < \beta$, $k \geq 0$, and thus it follows from (96) that

$$\sum_{k=0}^r \|\varepsilon(k)\|_2^2 < \beta^2 \|E(0)\|_F^2. \quad (97)$$

Letting $r \rightarrow \infty$, (97) implies that $\sum_{k=0}^{\infty} \|\varepsilon(k)\|_2^2 < \infty$, and thus we obtain (63). \square

REFERENCES

- [1] P. A. Nelson and S. J. Elliot, *Active Control of Sound*. New York: Academic, 1992.
- [2] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems*. New York: Wiley, 1996.
- [3] L. A. Sievers and A. H. von Flotow, "Comparison and extensions of control methods for narrow band disturbance rejection," *IEEE Trans. Signal Processing*, vol. 40, pp. 2377–2391, 1992.
- [4] W. Messner and M. Bodson, "Design of adaptive feedforward algorithms using internal model equivalence," *Int. J. Adaptive Contr. Signal Processing*, vol. 9, pp. 199–212, 1995.
- [5] S. J. Elliot, I. M. Stothers, and P. A. Nelson, "A multiple error LMS algorithm and its applications to the active control of sound and vibration," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1423–1434, 1987.
- [6] B. Widrow and E. Walach, *Adaptive Inverse Control*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [7] C. R. Fuller, C. A. Rogers, and H. H. Robertshaw, "Control of sound radiation with active/adaptive structures," *J. Sound Vibration*, vol. 157, pp. 19–39, 1992.
- [8] P. L. Feintuch, N. J. Bershad, and A. K. Lo, "A frequency domain model for 'filtered' LMS algorithm—Stability analysis, design, and elimination of the training mode," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-41, pp. 1518–1531, 1993.
- [9] F. Jiang, H. Tsuji, N. Ojima, H. Ohmori, and A. Sano, "Adaptation for active noise control," *IEEE Contr. Syst. Mag.*, pp. 36–47, Dec. 1997.
- [10] K. J. Astrom and B. Wittenmark, *Adaptive Control*, 2nd ed. Reading, MA: Addison-Wesley, 1995.
- [11] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [12] P. P. Kanjilal, *Adaptive Prediction and Predictive Control*. Herts, U.K.: Peter Peregrinus, 1995.
- [13] D. W. Clarke, C. H. Mohtadi, and P. S. Tuffs, "Generalized predictive algorithm—Part I, the basic algorithm," *Automatica*, vol. 23, pp. 137–148, 1987.
- [14] —, "Generalized predictive algorithm—Part II, extensions and interpretations," *Automatica*, vol. 23, pp. 149–160, 1987.

- [15] G. E. Garcia, D. M. Prett, and M. Morari, "Model predictive control theory and practice—A survey," *Automatica*, vol. 25, no. 3, pp. 335–348, 1989.
- [16] J. Richalet, "Industrial applications of model-based predictive control," *Automatica*, vol. 29, no. 5, pp. 1251–1274, 1993.
- [17] K. Furuta and M. Wongsaisuwana, "Discrete-time LQG dynamic controller design using plant Markov parameters," *Automatica*, vol. 31, pp. 1317–1324, 1995.
- [18] D. C. Hyland, E. G. Collins, Jr., W. M. Haddad, and D. L. Hunter, "Neural-network system identification for improved noise rejection," in *Proc. Amer. Contr. Conf.*, Seattle, WA, June 1995, pp. 345–349.
- [19] C. Ahn, "A recursive algorithm for identification of IIR systems," in *Proc. IEEE Conf. Dec. Contr.*, San Antonio, TX, Dec. 1993, pp. 2595–2600.
- [20] J. C. Akers and D. S. Bernstein, "Time-domain identification using ARMARKOV/Toeplitz models," in *Proc. Amer. Contr. Conf.*, Albuquerque, NM, June 1997, pp. 191–195.
- [21] —, "ARMARKOV least-squares identification," in *Proc. Amer. Contr. Conf.*, Albuquerque, NM, June 1997.
- [22] T. H. Van Pelt and D. S. Bernstein, "Least-squares identification using μ -Markov parameterizations," in *Proc. IEEE Conf. Decision Contr.*, Tampa, FL, Dec. 1998, pp. 618–619.
- [23] R. Venugopal and D. S. Bernstein, "Adaptive disturbance rejection using ARMARKOV/Toeplitz models," in *Proc. Amer. Contr. Conf.*, Albuquerque, NM, June 1997, pp. 1667–1661.
- [24] —, "Adaptive disturbance rejection using ARMARKOV system representations," in *Proc. IEEE Conf. Dec. Contr.*, San Diego, CA, Dec. 1997, pp. 1884–1889.
- [25] B. Francis, *A Course in H_∞ Control Theory*. New York: Springer-Verlag, 1987.
- [26] T. H. Van Pelt, R. Venugopal, and D. S. Bernstein, "Experimental comparison of adaptive noise control algorithms," in *Proc. Conf. Contr. Appl.*, Hartford, CT, Oct. 1997.



Ravinder Venugopal received the B. Tech degree in aerospace engineering from the Indian Institute of Technology, Madras, India, the M.S. degree in aerospace engineering from Texas A&M University, College Station, the M.S. degree in electrical engineering from the University of Michigan, Ann Arbor, and the Ph.D. degree in aerospace engineering from the University of Michigan in 1997.

From 1997 to 1999 he was a Postdoctoral Research Fellow at the Aerospace Engineering Department of the University of Michigan. He is currently an Applications Engineer with dSPACE, Inc. His research interests include discrete-time adaptive control, active noise and vibration control, and the development of control experiments for validating and motivating control theory.



Dennis S. Bernstein (M'82) received the Ph.D. degree from the University of Michigan, Ann Arbor, in 1982.

He has held positions in a government laboratory, industry, and academia. He is interested in virtually all applications of feedback control technology. Most of his current research is directed toward decreasing the dependence of modern control methods on plant and disturbance modeling to the greatest possible extent. His secondary interests include control education which he has written about in "A Student's Guide to Classical Control," *IEEE Control Systems*, vol. 17, pp. 96–100, Aug. 1997; "Control Experiments and What I Learned From Them: A Personal Journey," *IEEE Control Systems*, vol. 18, pp. 81–88, Apr. 1998; and "A Student's Guide to Research," *IEEE Control Systems*, vol. 19, pp. 102–108, Feb. 1999.