Estimation-Based Synthesis of H_∞-Optimal Adaptive FIR Filters for Filtered-LMS Problems

Bijan Sayyarrodsari, *Member, IEEE*, Jonathan P. How, Babak Hassibi, and Alain Carrier

Abstract—This paper presents a systematic synthesis procedure for H_∞-optimal adaptive FIR filters in the context of an active noise cancellation (ANC) problem. An estimation interpretation of the adaptive control problem is introduced first. Based on this interpretation, an H_∞ estimation problem is formulated, and its finite horizon prediction (filtering) solution is discussed. The solution minimizes the maximum energy gain from the disturbances to the predicted (filtered) estimation error and serves as the adaptation criterion for the weight vector in the adaptive FIR filter. We refer to this adaptation scheme as estimation-based adaptive filtering (EBAF). We show that the steady-state gain vector in the EBAF algorithm approaches that of the classical (normalized) filtered-X LMS algorithm. The error terms, however, are shown to be different. Thus, these classical algorithms can be considered to be approximations of our algorithm.

We examine the performance of the proposed EBAF algorithm (both experimentally and in simulation) in an active noise cancellation problem of a one-dimensional (1-D) acoustic duct for both narrowband and broadband cases. Comparisons to the results from a conventional filtered-LMS (FxLMS) algorithm show faster convergence without compromising steady-state performance and/or robustness of the algorithm to feedback contamination of the reference signal.

Index Terms—Adaptive filters, filtered LMS, optimal adaptive filters, robustness.

I. INTRODUCTION

LMS [1] has been the centerpiece of a wide variety of adaptive filtering algorithms for almost four decades. The straightforward derivation, and the simplicity of its implementation (especially at a time when computational power was a significant limitation) encouraged experiments with the algorithm in a diverse range of applications (see [1] and [2], for instance). In some applications, however, the simple implementation of LMS algorithm was found inadequate, and many attempts were made to overcome its shortcomings. A rich body of literature reflects the innovative solutions that have proved successful in practice. Commonly used algorithms such as normalized LMS, correlation LMS [3], leaky LMS [4], variable-step-size LMS [5], and filtered-LMS (FxLMS) [6] are

B. Sayyarrodsari is with Pavilion Technologies, Inc., Austin, TX 78758 USA (e-mail: bijan@pav.com).

J. P. How is with the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: jhow@mit.edu).

B. Hassibi is with the Mathematics and Sciences Department, Lucent Technologies, Murray Hill , NJ USA (e-mail: hassibi@research.bell-labs.com).

A. Carrier is with the Advanced Technology Center, Lockheed Martin, Palo Alto, CA, USA (e-mail: alain.carrier@lmco.com).

the outcome of such efforts. These algorithms employ the same adaptation criterion as the LMS algorithm, i.e., they use the instantaneous squared error to estimate the mean-square error [1]) and often assume slow adaptation to allow for the necessary linear operations in their derivation (see [2, Chs. 2 and 3], for instance). As [7] points out: "Many of the algorithms and approaches used are of an ad hoc nature; the tools are gathered from a wide range of fields; and good systematic approaches are still lacking." Introducing a systematic procedure for the synthesis of adaptive filters is one of the main goals of this paper.

Parallel to the efforts on the practical application of the LMS-based adaptive schemes, there has been a concerted effort to analyze these algorithms (see [8]–[14], for instance).[1] As the discussions in the above-mentioned references indicate, the analysis of an LMS-based adaptive algorithm often relies on simplifying assumptions (such as slow convergence of the adaptive filter, negligible contribution of the nonlinear/time-varying component of signals) and generally addresses only one particular aspect of its behavior. Narendra and Annaswamy [16] explain why this has been the case: ". . . adaptive systems are special classes of nonlinear systems . . . general methods for the analysis and synthesis of nonlinear systems do not exist since conditions for their stability can be established only on a system by system basis."

In this paper, we introduce a new framework for the synthesis and analysis of adaptive FIR filters for problems containing both a primary and secondary path. Whereas the connection between estimation problems and adaptive filtering problems devoid of a secondary path is well recognized (see, e.g., [22]), in this paper, we establish such a connection for problems containing a secondary path. In effect, we will show that an estimation interpretation of the "nonlinear" adaptive control problem leads to a carefully constructed "linear" estimation problem. The prediction (filtering) solution to this estimation problem then serves as the basis for an adaptation scheme for the weight vector of the adaptive FIR filter. Systematic analysis of the proposed algorithm can also be addressed in this framework using the analysis tools that are well developed in robust estimation.

This new framework is based on the recent results in robust estimation and is in the general spirit of [22]. Following the pioneering work in [17], robust control theory produced solutions [18], [19] that were designed to meet some performance criterion in the face of the limited knowledge of the exogenous dis-

turbances and imperfect system models. Further work in robust control and estimation (see [20] and [21] and references therein) produced straightforward solutions that allowed in depth study of the properties of the robust controllers/estimators. Of crucial importance for the work in this paper is the result in [22], where the H_∞-optimality of the LMS algorithm was established. Note that despite a long history of successful applications, prior to the work in [22], the LMS algorithm was regarded as an approximate solution to the least-squares minimization problem. The result in [22] proved that the LMS algorithm is indeed an exact solution to a minmax estimation problem and revealed a fundamental connection between an adaptive control algorithm (the LMS in this case) and a robust estimation problem.

Inspired by the analysis in [22], in this paper, we introduce an estimation interpretation of a general adaptive filtering problem (Fig. 1) and develop a systematic procedure for the synthesis of adaptive FIR filters based on this interpretation. The results of this paper demonstrate that the proposed estimation-based adaptive filtering (EBAF) algorithm is a systematic H_∞-based solution to the problem for which filtered-LMS algorithms have been developed in the past. In particular, it will be shown that the steady-state gain vector in the EBAF algorithm converges to the gain vector in the classical filtered-X LMS (normalized filtered-X LMS) adaptive algorithm when the prediction (filtering) solution to the equivalent estimation problem is used. The error term, however, is modified. Simulation results in this paper show that the Riccati-based gain vector in the EBAF algorithm leads to an improved transient behavior compared with its classical LMS-based counterparts. Improved steady-state performance of the new EBAF algorithm can be associated with the modification to the error term in the classical LMS-based adaptive algorithms.

The EBAF approach is applicable (due to its systematic nature) to both FIR and IIR adaptive filter design, but for simplicity, this paper only considers the FIR case. Furthermore, both the single and multichannel cases can be handled in this framework. We, however, will focus on the single channel scenario in this paper. Experimental results are included to validate the proposed adaptive algorithm. Matlab simulations are used to examine various aspects of the algorithm. Comparisons with a relevant, well-studied conventional adaptive algorithm (i.e., FxLMS) are also provided. We address the computational complexity of the proposed EBAF algorithm (see Section V) and show that it grows linearly with the length of the adaptive filter.

II. BACKGROUND

This section introduces the context in which the new estimation-based adaptive filtering (EBAF) algorithm will be presented. It defines the adaptive filtering problem of interest and describes the terminology that is used in this paper. We also outline a conventional solution to the problem based on the FxLMS algorithm. The discussion of key concepts of the EBAF algorithm and the mathematical formulation of the algorithm are left to Sections III and IV, respectively.

Referring to Fig. 1, the objective in this adaptive filtering problem is to adjust the weight vector in the adaptive FIR filter $W(k) = [w_0(k)\, w_1(k) \cdots w_N(k)]^T$ (k is the discrete time
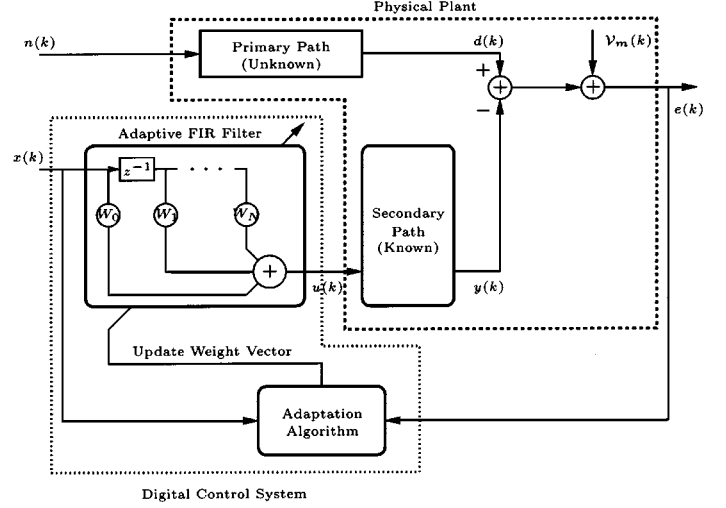

Fig. 1.   General block diagram for an active noise cancellation (ANC) problem.

index), such that the cancellation error $d(k) - y(k)$ is small in some appropriate measure. Note that $d(k)$ and $y(k)$ are outputs of the primary path $P(z)$ and the secondary path $S(z)$, respectively. Moreover

$n(k)$ input to the primary path;

$x(k)$ properly selected reference signal with a nonzero correlation with the primary input;

$u(k)$ control signal applied to the secondary path [generated as $u(k) \triangleq [x(k)\, x(k-1) \cdots x(k-N)]W(k)$];

$e(k)$ *measured* residual error available to the adaptation scheme.

Note that in a typical practice, $x(k)$ is obtained via some measurement of the primary input. The quality of this measurement will impact the correlation between the reference signal and the primary input. Similar to the conventional development of the FxLMS algorithm, however, this paper assumes perfect correlation between the two.

The FxLMS solution to this is shown in Fig. 2, where perfect correlation between the primary disturbance $n(k)$ and the reference signal $x(k)$ is assumed [1], [2]. Minimizing the *instantaneous* squared error $e^2(k)$ as an approximation to the mean-square error, FxLMS follows the LMS update criterion (i.e., to recursively adapt in the negative gradient direction)

$$W(k+1) = W(k) - \frac{\mu}{2}\nabla e^2(k)$$
$$e(k) = d(k) - y(k) = d(k) - \mathcal{S}(k) \oplus u(k)$$

where

μ adaptation rate;

$\mathcal{S}(k)$ impulse response of the secondary path;

"$\oplus$"   convolution.

Assuming slow adaptation, the FxLMS algorithm then *approximates* the instantaneous gradient in the weight vector update with

$$\nabla e^2(k) \cong -2\left[x'(k)\, x'(k-1) \cdots x'(k-N)\right]^T e(k)$$
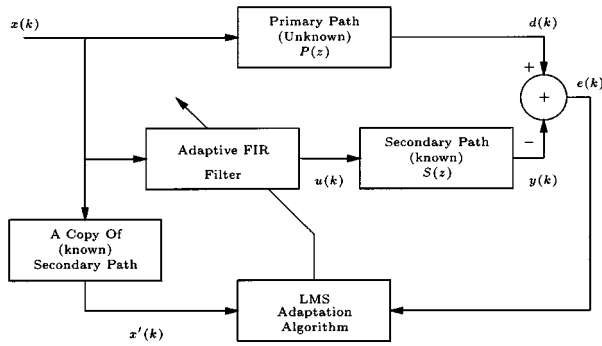$$\triangleq -2h'(k)e(k) \tag{1}$$

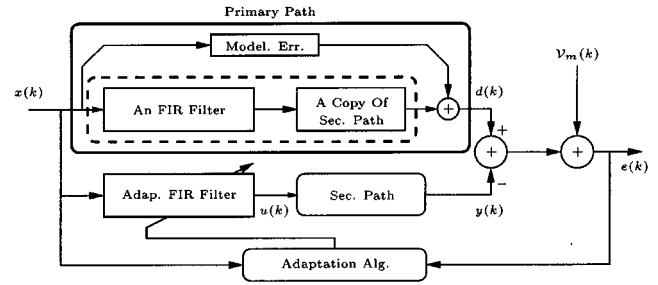Fig. 2.   Standard implementation of FxLMS algorithm.

Fig. 3.   Pictorial representation of the estimation interpretation of the adaptive control problem. Primary path is replaced by its approximate model.

where $x'(k) \triangleq \mathcal{S}(k) \oplus x(k)$ represents a filtered version of the reference signal that is available to the LMS adaptation [hence, the name (normalized) filtered-X LMS]. This yields the following adaptation criterion for the FxLMS algorithm:

$$W(k+1) = W(k) + \mu h'(k)e(k). \qquad (2)$$

In practice, however, only an approximate model of the secondary path (obtained via some identification scheme) is known, and it is this approximate model that is used to filter the reference signal.

A closely related adaptive algorithm is the one in which the adaptation rate is normalized with the estimate of the power of the reference vector, i.e.,

$$W(k+1) = W(k) + \mu \frac{h'(k)}{1 + \mu h^{*'}(k)h'(k)} e(k) \qquad (3)$$

where $*$ indicates complex conjugate. This algorithm is known as the normalized-FxLMS algorithm. For further discussion on the derivation and analysis of the FxLMS algorithm see [2] and [14].

III. EBAF ALGORITHM—MAIN CONCEPT

The principal goal of this section is to introduce the underlying concepts of the new EBAF algorithm. For the developments in this section, we assume perfect correlation between $n(k)$ and $x(k)$ in Fig. 1 [i.e., $x(k) = n(k)$ for all k]. This is the same condition under which the FxLMS algorithm was developed. We assume that the dynamics of the secondary path are known (e.g. by system identification). No explicit model for the primary path is needed.

As stated before, the objective in the adaptive filtering problem of Fig. 1 is to generate a control signal $u(k)$ such that the output of the secondary path $y(k)$ is "close" to the output of the primary path $d(k)$. To achieve this goal, for the given reference signal $x(k)$, the series connection of the FIR filter and the secondary path must constitute an appropriate model for the unknown primary path. In other words, with the adaptive FIR filter properly adjusted, the path from $x(k)$ to $d(k)$ must be equivalent to the path from $x(k)$ to $y(k)$. Based on this observation, in Fig. 3, the structure of the path from $x(k)$ to $y(k)$ is used to model the primary path. The modeling error is included to account for the imperfect cancellation. The above-mentioned observation forms the basis for the following

steps in an *estimation interpretation* of the adaptive control problem:

1) Introduce an approximate model for the primary path based on the architecture of the adaptive path from $x(k)$ to $y(k)$ (as shown in Fig. 3). There is an *optimal* value for the weight vector in the approximate model's FIR filter for which the modeling error is the smallest. This optimal weight vector, however, is not known. State-space models are used for both FIR filter and the secondary path.
2) In the approximate model for the primary path, use the available information to formulate an estimation problem that recursively estimates this optimal weight vector.
3) Adjust the weight vector of the adaptive FIR filter to the best available estimate of the optimal weight vector.

Before formalizing this estimation-based approach, we take a closer look at the signals (i.e., information) involved in Fig. 1. Note that $e(k) = d(k) - y(k) + \mathcal{V}_m(k)$, where

$e(k)$ available measurement;
$\mathcal{V}_m(k)$ exogenous disturbance that captures the effect of measurement noise, modeling error, and the initial condition uncertainty in error measurements;
$y(k)$ output of the secondary path;
$d(k)$ output of the primary path.

Note that unlike $e(k)$, the signals $y(k)$ and $d(k)$ are not directly measurable. With $u(k)$ fully known, however, the assumption of a known initial condition for the secondary path leads to the exact knowledge of $y(k)$. We later relax this assumption and consider the effect of an "inexact" initial condition in the performance of the adaptive filter (see Section VIII). We can now introduce the *derived* measured quantity that will be used in the estimation process

$$m(k) \triangleq e(k) + y(k) = d(k) + \mathcal{V}_m(k). \qquad (4)$$

IV. PROBLEM FORMULATION

Fig. 4 shows a block diagram representation of the approximate model to the primary path. We assume a state space model $[A_s(k), B_s(k), C_s(k), D_s(k)]$ for the secondary path. Note that both primary and secondary paths are assumed stable. We treat the weight vector $W(k) = [w_0(k)\, w_1(k) \cdots w_N(k)]^T$ as the state vector capturing the trivial dynamics $W(k+1) = W(k)$ that we assume for the FIR filter. With $\Theta(k)$ the state variable for the secondary path, $\Xi^T = (W^T(k)\, \Theta^T(k))$ is then the state vector
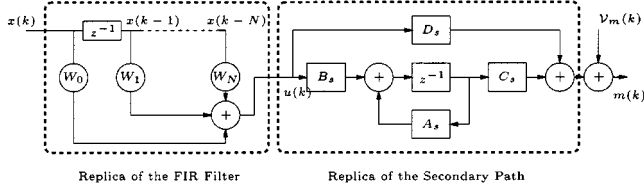
Fig. 4. Block diagram for the approximate model of the primary path.

for the overall system. The state space representation of the system is then

$$\begin{bmatrix} W(k+1) \\ \Theta(k+1) \end{bmatrix} = \begin{bmatrix} I_{(N+1)\times(N+1)} & 0 \\ B_s(k)h^*(k) & A_s(k) \end{bmatrix} \begin{bmatrix} W(k) \\ \Theta(k) \end{bmatrix} \triangleq F_k \Xi_k \tag{5}$$

where $h(k) = [x(k) \, x(k-1) \, \cdots \, x(k-N)]^T$ captures the effect of the reference input $x(\cdot)$. For this system, the *derived* measured output defined in (4) is

$$m(k) = [D_s(k)h^*(k) \quad C_s(k)] \begin{bmatrix} W(k) \\ \Theta(k) \end{bmatrix} + \mathcal{V}_m(k)$$
$$\triangleq H_k \Xi_k + \mathcal{V}_m(k). \tag{6}$$

We also define a linear combination of the states as the desired quantity to be estimated

$$s(k) = [L_{1,k} \quad L_{2,k}] \begin{bmatrix} W(k) \\ \Theta(k) \end{bmatrix} \triangleq L_k \Xi_k. \tag{7}$$

For simplicity, the single-channel problem is considered here, i.e., $m(k) \in \mathcal{R}^{1\times 1}$, $s(k) \in \mathcal{R}^{1\times 1}$, $\Theta(k) \in \mathcal{R}^{N_s \times 1}$, and $W(k) \in \mathcal{R}^{(N+1)\times 1}$. Extension to the multichannel case is straightforward and is discussed in [15]. There are several alternatives for selecting L_k and, thus, the variable to be estimated $s(k)$. The end goal of the estimation-based approach, however, is to set the weight vector in the adaptive FIR filter such that the output of the secondary path [$y(k)$ in Fig. 3] best matches $d(k)$. Therefore, we pick $s(k) = d(k)$, i.e., $L_k = H_k$. Any estimation algorithm can now be used to generate an estimate of the desired quantity $\hat{s}(k)$. Two main estimation approaches are considered next.

A. H_2 Optimal Estimation

Here, stochastic interpretation of the estimation problem is possible. Assuming that Ξ_0 (the initial condition for the system in Fig. 4) and $\mathcal{V}_m(\cdot)$ are zero mean uncorrelated random variables with known covariance matrices

$$E \begin{bmatrix} \Xi_0 \\ \mathcal{V}_m(k) \end{bmatrix} [\Xi_0^* \quad \mathcal{V}_m^*(j)] = \begin{bmatrix} \Pi_0 & 0 \\ 0 & Q_k \delta_{kj} \end{bmatrix} \tag{8}$$

$\hat{s}(k|k) \triangleq \mathcal{F}(m(0), \cdots, m(k))$, which is the causal linear least-mean-squares estimate of $s(k)$, is given by the Kalman filter recursions [23].

There are two primary difficulties with the H_2 optimal solution: a) The H_2 solution is optimal only if the stochastic assumptions are valid. If the external disturbance is not Gaussian (for instance when there is a considerable modeling error that should be treated as a component of the measurement disturbance), then

pursuing an H_2 filtering solution may yield undesirable performance; b) regardless of the choice for L_k, the recursive H_2 filtering solution does not simplify to the same extent as the H_∞ solution considered below. This can be of practical importance when the real-time computational power is limited. We therefore do not pursue the H_2 solution in this paper.

B. H_∞ Optimal Estimation

To avoid difficulties associated with the H_2 estimation, we consider a minmax formulation of the estimation problem in this section. This was first done in an adaptive filtering context in [22]. Here, the main objective is to limit the worst-case energy gain from the measurement disturbance and the initial condition uncertainty to the error in a causal (or strictly causal) estimate of $s(k)$. More specifically, we are interested in the following two cases. Let $\hat{s}(k|k) = \mathcal{F}_f(m(0), \cdots, m(k))$ denote an estimate of $s(k)$ given observations $m(i)$ for time $i = 0$ up to and including time $i = k$, and let $\hat{s}(k) \triangleq \hat{s}(k|k-1) = \mathcal{F}_p(m(0), \cdots, m(k-1))$ denote an estimate of $s(k)$ given $m(i)$ for time $i = 0$ up to and including $i = k-1$. Note that $\hat{s}(k|k)$ and $\hat{s}(k)$ are known as filtering and prediction estimates of $s(k)$, respectively. Two estimation errors can now be defined: the filtered error

$$e_{f,k} = \hat{s}(k|k) - s(k) \tag{9}$$

and the predicted error

$$e_{p,k} = \hat{s}(k) - s(k). \tag{10}$$

Given a final time M, the objective of the filtering problem can be formalized as finding $\hat{s}(k|k)$ such that for $\Pi_0 > 0$

$$\sup_{\mathcal{V}_m, \, \Xi_0} \frac{\displaystyle\sum_{k=0}^{M} e_{f,k}^* e_{f,k}}{(\Xi_0 - \hat{\Xi}_0)^* \Pi_0^{-1}(\Xi_0 - \hat{\Xi}_0) + \displaystyle\sum_{k=0}^{M} \mathcal{V}_m^*(k)\mathcal{V}_m(k)} \le \gamma^2 \tag{11}$$

for a given scalar $\gamma > 0$. Similarly, the objective of the prediction problem is to find $\hat{s}(k)$ *s.t.*

$$\sup_{\mathcal{V}_m, \, \Xi_0} \frac{\displaystyle\sum_{k=0}^{M} e_{p,k}^* e_{p,k}}{(\Xi_0 - \hat{\Xi}_0)^* \Pi_0^{-1}(\Xi_0 - \hat{\Xi}_0) + \displaystyle\sum_{k=0}^{M} \mathcal{V}_m^*(k)\mathcal{V}_m(k)} \le \gamma^2 \tag{12}$$

for a given scalar $\gamma > 0$. The question of optimality of the solution can be answered by finding the *infimum* value among all feasible γs. Note that for the H_∞ optimal estimation, there is no statistical assumption regarding the measurement disturbance. Therefore, the modeling error block (see Fig. 3) can be treated as a component of the measurement disturbance. For this reason, in Fig. 4, the "modeling error" block is eliminated.

V. H_∞-OPTIMAL SOLUTION

For the remainder of this paper, we consider the case where $L_k = H_k$. Note that this is the same assumption used in [22].

Referring to Fig. 4, this means that $s(k) = d(k)$. To discuss the solution, we draw from [23] the solutions to the γ-subop-timal finite-horizon filtering problem of (11) and the prediction problem of (12). Finally, we find the optimal value of γ and show how $\gamma = \gamma_{opt}$ simplifies the solutions.

A. γ-Suboptimal Finite Horizon Filtering Solution

Theorem 1 [23]: Consider the state space representation of the block diagram of Fig. 4, which is described by (5)–(7). A level-γ H_∞ filter that achieves (11) exists if, and only if, the matrices

$$R_k = \begin{bmatrix} I_p & 0 \\ 0 & -\gamma^2 I_q \end{bmatrix}$$

and

$$R_{e,k} = \begin{bmatrix} I_p & 0 \\ 0 & -\gamma^2 I_q \end{bmatrix} + \begin{bmatrix} H_k \\ L_k \end{bmatrix} P_k \begin{bmatrix} H_k^* & L_k^* \end{bmatrix} \quad (13)$$

(here, p and q are used to indicate the correct dimensions) have the same inertia[2] for all $0 \le k \le M$, where $P_0 = \Pi_0 > 0$ satisfies the Riccati recursion

$$P_{k+1} = F_k P_k F_k^* - K_{f,k} R_{e,k} K_{f,k}^* \quad (14)$$

where

$$K_{f,k} = (F_k P_k \begin{bmatrix} H_k^* & L_k^* \end{bmatrix}) R_{e,k}^{-1}. \quad (15)$$

If this is the case, then the central H_∞ estimator is given by

$$\hat{\Xi}_{k+1} = F_k \hat{\Xi}_k + K_{f,k} \left(m(k) - H_k \hat{\Xi}_k \right), \quad \hat{\Xi}_0 = 0 \quad (16)$$

$$\hat{s}(k|k) = L_k \hat{\Xi}_k + (L_k P_k H_k^*) R_{He,k}^{-1} \left(m(k) - H_k \hat{\Xi}_k \right) \quad (17)$$

with $K_{f,k} = (F_k P_k H_k^*) R_{He,k}^{-1}$ and $R_{He,k} = I_p + H_k P_k H_k^*$.
Proof: See [23].

B. γ-Suboptimal Finite Horizon Prediction Solution

Theorem 2 [23]: For the system described by (5)–(7), the level-γ H_∞ filter that achieves (12) exists if, and only if, all leading submatrices of

$$R_k^p = \begin{bmatrix} -\gamma^2 I_p & 0 \\ 0 & I_q \end{bmatrix}$$

and

$$R_{e,k}^p = \begin{bmatrix} -\gamma^2 I_p & 0 \\ 0 & I_q \end{bmatrix} + \begin{bmatrix} L_k \\ H_k \end{bmatrix} P_k \begin{bmatrix} L_k^* & H_k^* \end{bmatrix} \quad (18)$$

have the same inertia for all $0 \le k < M$. Note that P_k is updated according to (14). If this is the case, then one possible level-γ H_∞ filter is given by

$$\hat{\Xi}_{k+1} = F_k \hat{\Xi}_k + K_{p,k} \left(m(k) - H_k \hat{\Xi}_k \right), \quad \hat{\Xi}_0 = 0 \quad (19)$$

$$\hat{s}(k) = L_k \hat{\Xi}_k \quad (20)$$

$$K_{p,k} = F_k \tilde{P}_k H_k^* \left(I + H_k \tilde{P}_k H_k^* \right)^{-1} \quad (21)$$

$$\tilde{P}_k = \left(I - \gamma^{-2} P_k L_k^* L_k \right)^{-1} P_k. \quad (22)$$

Proof: See [23].

[2]By the inertia of a Hermitian matrix, we mean the number of its positive, negative, and zero eigenvalues.

Note that the condition in (18) is equivalent to

$$\left(I - \gamma^{-2} P_k L_k^* L_k \right) > 0, \quad \text{for} \quad k = 0, \cdots, M \quad (23)$$

and hence, $\tilde{P}_k$ in (22) is well defined. $\tilde{P}_k$ can also be defined as

$$\tilde{P}_k^{-1} = P_k^{-1} - \gamma^{-2} L_k^* L_k, \quad \text{for} \quad k = 0, \cdots, M \quad (24)$$

which proves useful in rewriting the prediction coefficient $K_{p,k}$ in (21), as follows. First, note that

$$F_k \tilde{P}_k H_k^* \left(I + H_k \tilde{P}_k H_k^* \right)^{-1} = F_k \left(\tilde{P}_k^{-1} + H_k^* H_k \right)^{-1} H_k^* \quad (25)$$

and hence, replacing for $\tilde{P}_k^{-1}$ from (24)

$$K_{p,k} = F_k \left(P_k^{-1} - \gamma^{-2} L_k^* L_k + H_k^* H_k \right)^{-1} H_k^*. \quad (26)$$

Theorems 1 and 2 (see Sections V-A and B) provide the form of the filtering and prediction estimators, respectively. The following section investigates the optimal value of γ for both of these solutions and outlines the simplifications that follow.

C. Optimal Value of γ

The optimal value of γ for the filtering solution will be discussed first. The discussion of the optimal prediction solution utilizes the results in the filtering case. The optimality discussions that follow are extensions to the results in [22] in which H_∞ optimality of the LMS algorithm for $\gamma = 1$ was derived.

1) Filtering Case:

a) $\gamma_{opt} \le 1$: We first show that for the filtering solution, $\gamma_{opt} \le 1$. Using (11), we can always pick $\hat{s}(k|k)$ to be simply $m(k)$. With this choice

$$\hat{s}(k|k) - s(k) = \mathcal{V}_m(k), \quad \text{for all} \quad k \quad (27)$$

and (11) reduces to

$$\sup_{\mathcal{V}_m \in \mathcal{L}_2, \, \Xi_0} \frac{\sum_{k=0}^{M} \mathcal{V}_m(k)^* \mathcal{V}_m(k)}{(\Xi_0 - \hat{\Xi}_0)^* \Pi_0^{-1} (\Xi_0 - \hat{\Xi}_0) + \sum_{k=0}^{M} \mathcal{V}_m(k)^* \mathcal{V}_m(k)} \quad (28)$$

which can never exceed 1 (i.e., $\gamma_{opt} \le 1$). We are therefore guaranteed a *feasible* solution for the H_∞ estimation problem in (11) when γ is chosen to be 1. Note that it is possible to directly demonstrate the feasibility of $\gamma = 1$. Using simple matrix manipulation, we can show that for $L_k = H_k$ and for $\gamma = 1$, R_k and $R_{e,k}$ have the same inertia for all k.

b) $\gamma_{opt} \ge 1$: To show that γ_{opt} is indeed 1, we must construct an admissible sequence of disturbances and a valid initial condition for which γ could be made arbitrarily close to 1 regardless of the filtering solution chosen. The necessary and sufficient conditions for the optimality of $\gamma_{opt} = 1$ are developed in the course of constructing this admissible sequence of disturbances.

Assume that $\hat{\Xi}_0^T = (\hat{W}_0^T \quad \hat{\Theta}_0^T)$ is the best estimate for the initial condition of the system in the approximate model of the primary path (see Fig. 4). Moreover, assume that $\hat{\Theta}_0$ is indeed the actual initial condition for the secondary path in Fig. 4. The

actual initial condition for the weight vector of the FIR filter in this approximate model is W_0. Then

$$m(0) = [\, D_s(0)h^*(0) \quad C_s(0)\,] \begin{bmatrix} W_0 \\ \hat{\Theta}_0 \end{bmatrix} + \mathcal{V}_m(0) \quad (29)$$

$$H_0\hat{\Xi}_0 = [\, D_s(0)h^*(0) \quad C_s(0)\,] \begin{bmatrix} \hat{W}_0 \\ \hat{\Theta}_0 \end{bmatrix} \quad (30)$$

where $m(0)$ is the (derived) measurement at time $k = 0$. Now, if we set

$$\mathcal{V}_m(0) = D_s(0)h^*(0)\left(\hat{W}_0 - W_0\right) = \mathcal{K}_\mathcal{V}(0)\left(\hat{W}_0 - W_0\right) \quad (31)$$

then $m(0) - H_0\hat{\Xi}_0 = 0$, and the estimate of the weight vector will not change. More specifically, (16) and (17) reduce to the following simple updates:

$$\hat{\Xi}_1 = F_0\hat{\Xi}_0, \quad \text{and} \quad \hat{s}(0|0) = L_0\hat{\Xi}_0 \quad (32)$$

which, given $L_0 = H_0$, generates the estimation error

$$e_{f,0} = \hat{s}(0|0) - s(0) = L_0\hat{\Xi}_0 - L_0\Xi_0$$
$$= D_s(0)h^*(0)\left(\hat{W}_0 - W_0\right) = \mathcal{V}_m(0). \quad (33)$$

Repeating a similar argument at $k = 1$ and 2, it is easy to see that if we pick

$$\mathcal{V}_m(1) = [D_s(1)h^*(1) + C_s(1)B_s(0)h^*(0)]\left(\hat{W}_0 - W_0\right)$$
$$\triangleq \mathcal{K}_\mathcal{V}(1)\left(\hat{W}_0 - W_0\right)$$
$$\mathcal{V}_m(2) = [D_s(2)h^*(2) + C_s(2)B_s(1)h^*(1)$$
$$+ C_s(2)A_s(1)B_s(0)h^*(0)]\left(\hat{W}_0 - W_0\right)$$
$$\triangleq \mathcal{K}_\mathcal{V}(2)\left(\hat{W}_0 - W_0\right)$$

then

$$m(k) - H_k\hat{\Xi}_k = 0, \quad \text{for} \quad k = 1, 2. \quad (34)$$

Note that when (34) holds, and with $L_k = H_k$, (17) reduces to

$$\hat{s}(k|k) = L_k\hat{\Xi}_k = H_k\hat{\Xi}_k \quad (35)$$

and hence

$$e_{f,k} = \hat{s}(k|k) - s(k) = \hat{s}(k|k) - [m(k) - \mathcal{V}_m(k)]$$
$$= H_k\hat{\Xi}_k - [m(k) - \mathcal{V}_m(k)] = \left[H_k\hat{\Xi}_k - m(k)\right] + \mathcal{V}_m(k)$$
$$= \mathcal{V}_m(k) \quad \text{for} \quad k = 1, 2. \quad (36)$$

Continuing this process, we can define $\mathcal{K}_\mathcal{V}(k)$, for $0 \leq k \leq M$, as seen in (37), shown at the bottom of the page, such that $\mathcal{V}_m(k), \forall k$ is an admissible disturbance. In this case, (11) reduces to (38), shown at the bottom of the page. From (37), note that

$$\sum_{k=0}^{M} \mathcal{K}_\mathcal{V}^*(k)\mathcal{K}_\mathcal{V}(k) = \Lambda_M^* \Delta_M^* \Delta_M \Lambda_M = \|\Delta_M \Lambda_M\|_2^2 \quad (39)$$

and hence, the ratio in (38) can be made arbitrarily close to one if

$$\lim_{M \to \infty} \|\Delta_M \Lambda_M\|_2 \to \infty. \quad (40)$$

Equation (40) will be referred to as the *condition for optimality* of $\gamma = 1$ for the filtering solution. Equation (40) can now be used to derive necessary and sufficient conditions for optimality of $\gamma = 1$. First, note that a *necessary condition* for (40) is

$$\lim_{M \to \infty} \|\Lambda_M\|_2 \to \infty \quad (41)$$

$$\begin{bmatrix} \mathcal{K}_\mathcal{V}(0) \\ \mathcal{K}_\mathcal{V}(1) \\ \mathcal{K}_\mathcal{V}(2) \\ \vdots \\ \mathcal{K}_\mathcal{V}(M) \end{bmatrix} = \begin{bmatrix} D_s(0) & 0 & 0 & 0 & \cdots & 0 \\ C_s(1)B_s(0) & D_s(1) & 0 & 0 & \cdots & 0 \\ C_s(2)A_s(1)B_s(0) & C_s(2)B_s(1) & D_s(2) & 0 & \cdots & 0 \\ & & & \ddots & & \vdots \\ \vdots & & \cdots & & & D_s(M) \end{bmatrix} \begin{bmatrix} h(0) \\ h(1) \\ h(2) \\ \vdots \\ h(M) \end{bmatrix} \triangleq \Delta_M \Lambda_M \quad (37)$$

$$\sup_{\Xi_0} \frac{\displaystyle\sum_{k=0}^{M} \mathcal{V}_m(k)^*\mathcal{V}_m(k)}{(\Xi_0 - \hat{\Xi}_0)^*\Pi_0^{-1}(\Xi_0 - \hat{\Xi}_0) + \displaystyle\sum_{k=0}^{M} \mathcal{V}_m(k)^*\mathcal{V}_m(k)}$$
$$= \sup_{\Xi_0} \frac{(\hat{W}_0 - W_0)^* \left[\displaystyle\sum_{k=0}^{M} \mathcal{K}_\mathcal{V}^*(k)\mathcal{K}_\mathcal{V}(k)\right](\hat{W}_0 - W_0)}{(\Xi_0 - \hat{\Xi}_0)^*\Pi_0^{-1}(\Xi_0 - \hat{\Xi}_0) + (\hat{W}_0 - W_0)^* \left[\displaystyle\sum_{k=0}^{M} \mathcal{K}_\mathcal{V}^*(k)\mathcal{K}_\mathcal{V}(k)\right](\hat{W}_0 - W_0)} \quad (38)$$

or equivalently

$$\lim_{M\to\infty} \sum_{k=0}^{M} h^*(k)h(k) \;\to\; \infty. \tag{42}$$

The $h(k)$ that satisfies the condition in (42) is referred to as *exciting* [22]. Several sufficient conditions can now be developed. Since

$$\|\Delta_M \Lambda_M\|_2 \geq \sigma_{\min}(\Delta_M)\|\Lambda_M\|_2 \tag{43}$$

one *sufficient condition* is that

$$\sigma_{\min}(\Delta_M) > \epsilon, \quad \forall M, \quad \text{and} \quad \epsilon > 0. \tag{44}$$

For LTI systems, the sufficient condition (44) is equivalent to the system having no zeros on the unit circle. Another sufficient condition is that $h(k)$s be *persistently exciting*, i.e.,

$$\lim_{M\to\infty} \sigma_{\min}\left[\frac{1}{M}\sum_{k=0}^{M} h(k)h^*(k)\right] > 0 \tag{45}$$

which holds for most reasonable systems.

2) Prediction Case: The optimal value for γ cannot be less than one in the prediction case. In the previous section, we showed that despite using all available measurements up to and including time k, the sequence of the admissible disturbances $\mathcal{V}_m(k) = \mathcal{K}_\mathcal{V}(k)(\hat{W}_0 - W_0)$ for $k = 0, \cdots, M$ [where $\mathcal{K}_\mathcal{V}(k)$ is given by (37)] prevented the filtering solution from achieving $\gamma < 1$. The prediction solution that uses only the measurements up to time k (not including k itself) cannot improve over the filtering solution, and therefore, the energy gain γ is at least one.

Next, we show that if the initial condition P_0 is chosen appropriately (i.e., if it is small enough), then $\gamma_{opt} = 1$ can be guaranteed. Referring to the Lyapunov recursion of (61), the Riccati matrix at time k can be written as

$$P_k = \left(\prod_{j=0}^{k-1} F_j\right) P_0 \left(\prod_{j=0}^{k-1} F_j\right)^*$$
$$F_j = \begin{bmatrix} I & 0 \\ B_s(j)h^*(j) & A_s(j) \end{bmatrix}. \tag{46}$$

Defining

$$\Psi_A^j = A_s(j)A_s(j-1)\cdots A_s(0) \tag{47}$$

(46) can be written as

$$P_k = \begin{bmatrix} I & 0 \\ \sum_{j=0}^{k-1}\Psi_A^j B_s(j)h^*(k-1-j) & \Psi_A^k \end{bmatrix}$$
$$\cdot P_0 \begin{bmatrix} I & 0 \\ \sum_{j=0}^{k-1}\Psi_A^j B_s(j)h^*(k-1-j) & \Psi_A^k \end{bmatrix}^*. \tag{48}$$

From Theorem 2 in Section V-B, the condition for the existence of a prediction solution is $(I - \gamma^{-2}P_k L_k^* L_k) > 0$, or equivalently

$$(\gamma^2 - L_k P_k L_k^*) > 0. \tag{49}$$

Note that $L_k = [D_s(k)h^*(k) \; C_c(k)]$, and therefore, (49) can be rewritten as

$$\gamma^2 - [D_s(k)h^*(k) \quad C_s(k)] P_k \begin{bmatrix} h(k)D_s^*(k) \\ C_s^*(k) \end{bmatrix} > 0. \tag{50}$$

Replacing for P_k from (38) and carrying out the matrix multiplications, (50) yields

$$\gamma^2 - \left[\begin{array}{c} h(k)D_s^*(k) + \sum_{j=0}^{k-1} h(k-1-j)B_s^*(j)\Psi_A^{*j}C_s^*(k) \\ \Psi_A^{*k}C_s^*(k) \end{array} \right]^*$$
$$\cdot P_0 \left[\begin{array}{c} h(k)D_s^*(k) + \sum_{j=0}^{k-1} h(k-1-j)B_s^*(j)\Psi_A^{*j}C_s^*(k) \\ \Psi_A^{*k}C_s^*(k) \end{array} \right] > 0. \tag{51}$$

Introducing

$$h'^*(k) = D_s h^*(k) + \sum_{j=0}^{k-1} C_s(k)\Psi_A^j B_s(j)h^*(k-1-j) \tag{52}$$

as the filtered version of the reference vector $h(k)$, we can express (51) as

$$\gamma^2 - [h'^*(k) \quad C_s(k)\Psi_A^k] P_0 \begin{bmatrix} h'(k) \\ \Psi_A^{*k}C_s^*(k) \end{bmatrix} > 0. \tag{53}$$

Selecting the initial value of the Riccati matrix, without loss of generality, as

$$P_0 = \begin{bmatrix} \mu I & 0 \\ 0 & \alpha I \end{bmatrix} \tag{54}$$

and (53) reduces to

$$\gamma^2 - \mu h'^*(k)h'(k) - \alpha C_s(k)\Psi_A^k \Psi_A^{*k} C_s^*(k) > 0. \tag{55}$$

It is now clear that a prediction solution for $\gamma = 1$ exists if

$$\mu < \frac{1 - \alpha C_s(k)\Psi_A^k \Psi_A^{*k} C_s^*(k)}{h'^*(k)h'(k)}. \tag{56}$$

Equation (56) is therefore the condition for optimality of $\gamma_{opt} = 1$ for the prediction solution.

D. Simplified Solution Due to $\gamma = 1$

1) Filtering Case: Now, we show that with $H_k = L_k$ and $\gamma = 1$, the Riccati equation (14) is considerably simplified. Applying the matrix inversion lemma to $R_{e,k}$, it can be shown that [15]

$$R_{e,k}^{-1} = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} - \begin{bmatrix} H_k \\ -H_k \end{bmatrix} P_k [H_k^* \quad -H_k^*]. \tag{57}$$

Therefore

$$K_{f,k}R_{e,k}K_{f,k}^*$$
$$= (F_k P_k [H_k^* \quad H_k^*]) R_{e,k}^{-1} \left(\begin{bmatrix} H_k \\ H_k \end{bmatrix} P_k F_k^*\right) = 0$$

for $\gamma = 1$ and for all k. Thus, the Riccati recursion (14) reduces to the *Lyapunov* recursion $P_{k+1} = F_k P_k F_k^*$ with $P_0 = \Pi_0 > 0$.

Partitioning the Riccati matrix P_k in block matrices conformable with the block matrix structure of F_k, (14) yields the following simple update:

$$\begin{cases} P_{11, k+1} = P_{11, k}, \quad P_{11, 0} = \Pi_{11, 0} \\ P_{12, k+1} = P_{12, k} A_s(k) + P_{11, k} h(k) B_s^*(k), \\ \quad P_{12, 0} = \Pi_{12, 0} \\ P_{22, k+1} = B_s(k) h(k)^* P_{11, k} h(k) B_s^*(k) \\ \quad + A_s(k) P_{12, k}^* h(k) B_s^*(k) \\ \quad + B_s(k) h^*(k) P_{12, k} A_s^*(k) \\ \quad + A_s(k) P_{22, k} A_s^*(k), \quad P_{22, 0} = \Pi_{22, 0}. \end{cases} \quad (58)$$

We can now summarize the filtering solution in the following theorem.

Theorem 3: Consider the system described by (5)–(7), with $L_k = H_k$. If the optimality condition (40) is satisfied, the H_∞-optimal filtering solution achieves $\gamma_{opt} = 1$, and the central H_∞-optimal filter is given by

$$\hat{\Xi}_{k+1} = F_k \hat{\Xi}_k + K_{f, k} \left(m(k) - H_k \hat{\Xi}_k \right), \quad \hat{\Xi}_0 = 0 \quad (59)$$

$$\hat{s}(k|k) = L_k \hat{\Xi}_k + (L_k P_k H_k^*) R_{He, k}^{-1} \left(m(k) - H_k \hat{\Xi}_k \right) \quad (60)$$

with $K_{f, k} = (F_k P_k H_k^*) R_{He, k}^{-1}$ and $R_{He, k} = I_p + H_k P_k H_k^*$, where P_k satisfies the Lyapunov recursion

$$P_{k+1} = F_k P_k F_k^*, \quad P_0 = \Pi_0. \quad (61)$$

Proof: The proof rfollows from the discussions above.

2) Prediction Case: Referring to (26), it is clear that for $\gamma = 1$ and for $L_k = H_k$, the coefficient $K_{p, k}$ will reduce to $F_k P_k H_k^*$. Therefore, we can summarize the prediction solution as follows.

Theorem 4: Consider the system described by (5)–(7), with $L_k = H_k$. If the optimality conditions (40) and (56) are satisfied, and with P_0 as defined in (54), the H_∞-optimal prediction solution achieves $\gamma_{opt} = 1$, and the central filter is given by

$$\hat{\Xi}_{k+1} = F_k \hat{\Xi}_k + K_{p, k} \left(m(k) - H_k \hat{\Xi}_k \right), \quad \hat{\Xi}_0 = 0 \quad (62)$$

$$\hat{s}(k) = L_k \hat{\Xi}_k \quad (63)$$

with $K_{p, k} = F_k P_k H_k^*$, where P_k satisfies the Lyapunov recursion (61).

Proof: The proof follows from the discussions above.

VI. IMPORTANT REMARKS

We start by the statement of the main idea in EBAF algorithm. At a given time k, use the available information on

  a) measurement history $e(i)$ for $0 \le i \le k$;
  b) control history $u(i)$ for $0 \le i < k$;
  c) reference signal history $x(i)$ for $0 \le i \le k$;
  d) the model of the secondary path and the estimate of its initial condition;
  e) the predetermined length of the adaptive FIR filter to produce the best estimate of the actual output of the primary path $d(k)$.

The key premise is that if we can accurately estimate $d(k)$, then we can generate the inputs $u(k)$ that cancel it. The objective of the EBAF algorithm is to make $y(k)$ match the optimal estimate of $d(k)$ (see Fig. 3). For the adaptive filtering problem in Fig. 1, however, we only have direct access to the weight vector of the adaptive FIR filter. Because of this practical constraint, the EBAF algorithm adapts the weight vector in the adaptive FIR filter according to the estimate of the optimal weight vector given by (59) or (62) (for the filtering, or prediction solutions, respectively). Note that $\hat{\Xi}_k^T = (\hat{W}^T(k) \, \hat{\Theta}^T(k))$. The error analysis for this adaptive algorithm is discussed in Section VIII. Now, we can highlight main features of this algorithm as follows:

1) The estimation-based adaptive filtering (EBAF) algorithm yields a solution that only requires one Riccati recursion. The recursion propagates *forward* in time and does not require any information about the future of the system or the reference signal (thus allowing the resulting adaptive algorithm to be real-time implementable). This has come at the expense of restricting the controller to an FIR structure in advance.

2) With $K_{f, k} R_{e, k} K_{f, k}^* = 0$, $P_{k+1} = F_k P_k F_k^*$ is the simplified Riccati equation, which considerably reduces the computational complexity involved in propagating the Riccati matrix. Furthermore, this Riccati update always generates a non-negative definite P_k, as long as P_0 is selected to be positive definite [see (61)].

3) In general, the solution to an H_∞ filtering problem requires verification of the fact that R_k and $R_{e, k}$ are of the same inertia at each step [see (13)]. In a similar way, the prediction solution requires that all sub-matrices of R_k^p and $R_{e, k}^p$ have the same inertia for all k [see (18)]. This can be a computationally expensive task. Moreover, it may lend to a breakdown in the solution if the condition is not met at some time k. Our formulation of the problem eliminates the need for such checks, as well as the potential breakdown of the solution, by providing a definitive answer to the feasibility and optimality of $\gamma = 1$.

4) When $[A_s(k), B_s(k), C_s(k), D_s(k)] = [0, 0, 0, I]$ for all k [i.e., the output of the FIR filter directly cancels $d(k)$ in Fig. 1], then the filtering/prediction results we have derived reduce to the simple normalized-LMS/LMS algorithms in [22], as expected.

5) As mentioned earlier, there is no need to verify the solutions at each time step; therefore, the computational complexity of the estimation-based approach is $O(n^3)$ (primarily for calculating $F_k P_k F_K^*$), where $n = (N + 1) + N_s$. Here, $(N + 1)$ is the length of the FIR filter, and N_s is the order of the secondary path. The special structure of F_k, however, reduces the computational complexity to $O(N_s^3 + N_s N)$, i.e., cubic in the order of the secondary path and linear in the length of the FIR filter [see (58)]. This is often a substantial reduction in the computation since $N_s \ll N$. Note that the computational complexity for FxLMS is quadratic in N_s and linear in N.

VII. IMPLEMENTATION SCHEME FOR EBAF ALGORITHM

Three sets of variables are used to describe the implementation scheme.

1) *Best Available Estimate of a Variable:* Referring to (16) and (19) and noting the fact that $\hat{\Xi}_k^T = (\hat{W}^T(k)\hat{\Theta}^T(k))$, we can define $\hat{W}(k)$ as the estimate of the weight vector and $\hat{\Theta}(k)$ as the secondary path state estimate in the approximate model of the primary path.

2) *Actual Value of a Variable:* Referring to Fig. 1, we define $u(k) \triangleq h^*(k)\hat{W}(k)$ as the actual input to the secondary path, $y(k)$ as the actual output of the secondary path, and $d(k)$ as the actual output of the primary path. Note that $d(k)$ and $y(k)$ are not directly measurable and that at each iteration, the weight vector in the adaptive FIR filter is set to $\hat{W}(k)$.

3) *Adaptive Algorithm's Internal Copy of a Variable:* Recall that in (4), we used $y(k)$ to construct the derived measurement $m(k)$. Since $y(k)$ is not directly available, the adaptive algorithm needs to generate an internal copy of this variable. This internal copy [referred to as $y_{copy}(k)$] is constructed by applying $u(k)$ (the actual control signal) to a model of the secondary path inside the adaptive algorithm. The initial condition for this model is $\Theta_{copy}(0)$. In other words, the derived measurement is constructed as follows:

$$\Theta_{copy}(k+1) = A_s(k)\Theta_{copy}(k) + B_s(k)u(k) \qquad (64)$$

$$y_{copy}(k) = C_s(k)\Theta_{copy}(k) + D_s(k)u(k) \qquad (65)$$

$$m(k) = e(k) + y_{copy}(k). \qquad (66)$$

Given the identified model for the secondary path and its input $u(k) = h^*(k)\hat{W}(k)$, the adaptive algorithm's copy of $y(k)$ will be exact if the actual initial condition of the secondary path is known. Obviously, one cannot expect to have the exact knowledge of the actual initial condition of the secondary path. In the next section, however, we show that when the secondary path is linear and stable, the contribution of the initial condition to its output decreases to zero as k increases. Therefore, the internal copy of $y(k)$ will converge to the actual value of $y(k)$ over time.

Now, we can outline the implementation algorithm as follows.

1) Start with $\hat{W}(0) = 0$ and $\hat{\Theta}(0) = 0$ as the initial guess for the state vector in the approximate model of the primary path. In addition. assume that $\Theta_{copy}(0) = 0$, and $h(0) = [x(0)\,0\,\cdots\,0]^T$. The initial value for the Riccati matrix is P_0, which is chosen to be block diagonal. The role of P_0 is similar to the learning rate in LMS-based adaptive algorithms (see Section V-C.2).

2) If $0 \le k \le M$ (finite horizon) do the following:

   a) Form the control signal

$$u(k) = h^*(k)\hat{W}(k) \qquad (67)$$

   to be applied to the secondary path. Note that applying $u(k)$ to the secondary path produces

$$y(k) = C_s(k)\Theta(k) + D_s(k)u(k) \qquad (68)$$

at the output of the secondary path. This in turn leads to the following error signal measured at time k:

$$e(k) = d(k) - y(k) + \mathcal{V}_m(k) \qquad (69)$$

which is available to the adaptive algorithm to perform the state update at time k.

   b) Propagate the state estimate and the internal copy of the state of the secondary path as follows:

$$\begin{bmatrix} \begin{bmatrix} \hat{W}(k+1) \\ \hat{\Theta}(k+1) \end{bmatrix} \\ \Theta_{copy}(k+1) \end{bmatrix}$$

$$= \begin{bmatrix} [F_k + K_{f,k}[0 - C_s(k)]] & K_{f,k}C_s(k) \\ [B_s(k)h^*(k) \quad 0] & A_s(k) \end{bmatrix} \begin{bmatrix} \begin{bmatrix} \hat{W}(k) \\ \hat{\Theta}(k) \end{bmatrix} \\ \Theta_{copy}(k) \end{bmatrix}$$

$$+ \begin{bmatrix} [K_{f,k}] \\ 0 \end{bmatrix} e(k) \qquad (70)$$

   where $e(k)$ is the error sensor measurement at time k given by (69), and $K_{f,k} = F_k P_k H_k^*(I + H_k P_k H_k^*)^{-1}$ (see Theorem 3). Note that for the prediction-based EBAF algorithm, we only need to replace $K_{f,k}$ with $K_{p,k} = F_k P_k H_k^*$.

   c) Update the Riccati matrix P_k using the Lyapunov recursion

$$\begin{bmatrix} P_{11} & P_{12,k+1} \\ P_{12,k+1}^* & P_{22,k+1} \end{bmatrix}$$
$$= \begin{bmatrix} I & 0 \\ B_s(k)h^*(k) & A_s(k) \end{bmatrix}$$
$$\cdot \begin{bmatrix} P_{11} & P_{12,k} \\ P_{12,k}^* & P_{22,k} \end{bmatrix} \begin{bmatrix} I & 0 \\ B_s(k)h^*(k) & A_s(k) \end{bmatrix}^* \qquad (71)$$

   P_{k+1} will be used in (70) to update the state estimate.

3) Go to 2.

VIII. ERROR ANALYSIS

In Section VII, we pointed out that the proposed implementation scheme can deviate from an H_∞-optimal solution for two main reasons: 1) the error in the initial condition of the secondary path that can cause y_{copy} to be different from $y(k)$ and 2) the additional error in the cancellation of $d(k)$ due to the fact that we cannot directly set $y(k)$ to be $\hat{s}(k|k)$ [or $\hat{s}(k)$]. All we can do is to set the weight vector in the adaptive FIR filter to be $\hat{W}(k)$.

Here, we will discuss both errors in detail.

A. Effect of Initial Condition

As our earlier discussions indicate, the secondary path in Fig. 1 is assumed to be linear. For a linear system, the output at any given time can be decomposed into two components: the *zero-input component*, which is associated with the portion of the output solely due to the initial condition of the system, and

the *zero-state component*, which is the portion of the output solely due to the input to the system.

For a stable system, the zero-input component of the response will decay to zero for large k. Therefore, any difference between $y_{copy}(k)$ and $y(k)$ (which with a known input to the secondary path can only be due to the unknown initial condition) will go to zero as k grows. In other words, exact knowledge of the initial condition of the secondary path does not affect the performance of the proposed EBAF algorithm for sufficiently large k.

B. Effect of Practical Limitation in Setting $y(k)$ to $\hat{s}(k|k)$ ($\hat{s}(k)$)

As pointed out earlier, the physical setting of the adaptive control problem in Fig. 1 only allows for the weight vector in the adaptive FIR filter to be adjusted to $\hat{W}(k)$. In other words, the state of the secondary path cannot be set to a desired value at each step. Instead, Θ_k evolves based on its initial condition and the control input $u(k)$ that we provide. Assume that $\Theta(k)$ is the actual state of the secondary path at time k. The actual output of the secondary path is then

$$y(k) = D_s(k)h^*(k)\hat{W}(k) + C_s(k)\Theta(k) \quad (72)$$

which leads to the following cancellation error:

$$d(k) - y(k) = d(k) - \left(D_s(k)h^*(k)\hat{W}(k) + C_s(k)\Theta(k)\right). \quad (73)$$

For the prediction solution of Theorem 4, if we add the zero quantity $\pm C_s(k)\hat{\Theta}(k)$ to the right-hand side of (73) and taking the norm of both sides, we can write

$$\|d(k) - y(k)\| = \|d(k) - (D_s(k)h^*(k)\hat{W}(k) + C_s(k)\Theta(k))$$
$$\pm C_s(k)\hat{\Theta}(k)\|$$
$$= \|(d(k) - D_s(k)h^*(k)\hat{W}(k) - C_s(k)\hat{\Theta}(k))$$
$$+ C_s(k)(\hat{\Theta}(k) - \Theta(k))\|.$$

Therefore

$$\frac{\|d(k) - y(k)\|}{\tilde{\Xi}_0^*\Pi_0^{-1}\tilde{\Xi}_0 + \sum_{k=0}^{M} \mathcal{V}_m^*(k)\mathcal{V}_m(k)}$$
$$\leq \frac{\|d(k) - \hat{s}(k)\|}{\tilde{\Xi}_0^*\Pi_0^{-1}\tilde{\Xi}_0 + \sum_{k=0}^{M} \mathcal{V}_m^*(k)\mathcal{V}_m(k)}$$
$$+ \frac{\|C_s(k)(\hat{\Theta}(k) - \Theta(k))\|}{\tilde{\Xi}_0^*\Pi_0^{-1}\tilde{\Xi}_0 + \sum_{k=0}^{M} \mathcal{V}_m^*(k)\mathcal{V}_m(k)} \quad (74)$$

where $\tilde{\Xi}_0 = (\Xi_0 - \hat{\Xi}_0)$ and Ξ_k is defined in (5). Note that the first term on the right-hand side of (74) is the prediction error energy gain [see (12)]. Therefore, the energy gain of the cancellation error with the prediction-based EBAF exceeds the error energy gain of the H_∞ optimal prediction solution by the second term on the right-hand side of (74). It can be shown that when the primary inputs $h(k)$ are persistently exciting [see (45)], the dynamics for the state estimation error $\hat{\Theta}(k) - \Theta(k)$

are internally stable, which implies that the second term on the right-hand side of (74) is bounded for all M and in the limit when $M \to \infty$.[3]

When $D_s(k) = 0$ for all k, an implementation of the filtering solution that utilizes the most recent measurement $m(k)$ is feasible. In this case, the filtering solution in (16) and (17) can be written as follows:

$$\check{\Xi}_{k|k} = \hat{\Xi}_k + P_k H_k^*(I_p + H_k P_k H_k^*)^{-1}\left(m(k) = H_k\hat{\Xi}_k\right) \quad (75)$$

and

$$\hat{\Xi}_{k+1} = F_k\check{\Xi}_{k|k} \quad (76)$$

$$\hat{s}(k|k) = L_k\check{\Xi}_{k|k} \quad (77)$$

where the weight vector update in the adaptive FIR filter follows (75). With a derivation identical to the one for prediction solution, it can be shown that the performance bound in this case is

$$\frac{\|d(k) - y(k)\|}{\tilde{\Xi}_0^*\Pi_0^{-1}\tilde{\Xi}_0 + \sum_{k=0}^{M} \mathcal{V}_m^*(k)\mathcal{V}_m(k)}$$
$$\leq \frac{\|d(k) - \hat{s}(k|k)\|}{\tilde{\Xi}_0^*\Pi_0^{-1}\tilde{\Xi}_0 + \sum_{k=0}^{M} \mathcal{V}_m^*(k)\mathcal{V}_m(k)}$$
$$+ \frac{\|C_s(k)(\hat{\Theta}(k|k) - \Theta(k))\|}{\tilde{\Xi}_0^*\Pi_0^{-1}\tilde{\Xi}_0 + \sum_{k=0}^{M} \mathcal{V}_m^*(k)\mathcal{V}_m(k)}. \quad (78)$$

An argument similar to the prediction case shows that the second term on the right-hand side has a finite gain as well.

IX. RELATIONSHIP TO THE NORMALIZED-FxLMS/FxLMS ALGORITHMS

In this section, we will first show that as $k \to \infty$, the gain vector in the prediction-based EBAF algorithm converges to the gain vector in the classical FxLMS algorithm. Thus, FxLMS is an approximation to the steady-state EBAF. The error terms in the two algorithms are shown to be different [compare (84) and (2)]. Therefore, it is expected that the prediction-based EBAF demonstrate superior transient performance compared with the FxLMS algorithm. Simulation results in the next section agree with this expectation. The fact that the gain vectors asymptotically coincide agrees with the fact that the derivation of the FxLMS algorithm relies on the assumption that the adaptive filter and the secondary path are interchangeable, which can only be true in the steady state. Similar results are shown for the connection between the filtering-based EBAF and the normalized FxLMS adaptive algorithms.

For the discussion in this section, the secondary path is assumed, for simplicity, to be LTI, i.e., $[A_s, B_s, C_s, D_s]$. Note

---

[3]Green and Limebeer [19] show that if the exogenous disturbance is assumed to be a zero mean white noise process with unit intensity and independent of the initial condition of the system Ξ_0, then the terminal state estimation error variance satisfies

$$\mathcal{E}(\Xi_k - \hat{\Xi}_k)(\Xi_k - \hat{\Xi}_k)^* \leq P_k.$$

that for the LTI system, Ψ_A^k in (47) reduces to A_s^k. The Riccati matrix P_k in (48) can then be rewritten as

$$
P_k = \begin{bmatrix} I & 0 \\ \sum_{j=0}^{k-1} A_s^j B_s h^*(k-1-j) & A_s^k \end{bmatrix}
$$
$$
\cdot P_0 \begin{bmatrix} I & 0 \\ \sum_{j=0}^{k-1} A_s^j B_s h^*(k-1-j) & A_s^k \end{bmatrix}^*. \quad (79)
$$

Equation (79) will be used in establishing the proper connections between the filtered/predicted solutions of Section V and the conventional normalized-FxLMS/FxLMS algorithms.

A. Prediction Solution and Its Connection to the FxLMS Algorithm

To study the asymptotic behavior of the state estimate update, note that for an stable secondary path $A_s^k \to 0$ as $k \to \infty$. Therefore, using (79)

$$
P_k \to \begin{bmatrix} I & 0 \\ \sum_{j=0}^{k-1} A_s^j B_s h^*(k-1-j) & 0 \end{bmatrix}
$$
$$
\cdot P_0 \begin{bmatrix} I & 0 \\ \sum_{j=0}^{k-1} A_s^j B_s h^*(k-1-j) & 0 \end{bmatrix}^* \text{ as } k \to \infty \quad (80)
$$

which for

$$
P_0 = \begin{bmatrix} P_{11}(0) & P_{12}(0) \\ P_{21}(0) & P_{22}(0) \end{bmatrix}
$$

results in

$$
P_k \to \begin{bmatrix} I \\ \sum_{j=0}^{k-1} A_s^j B_s h^*(k-1-j) \end{bmatrix}
$$
$$
\cdot P_{11}(0) \begin{bmatrix} I \\ \sum_{j=0}^{k-1} A_s^j B_s h^*(k-1-j) \end{bmatrix}^* \text{ as } k \to \infty. \quad (81)
$$

Selecting $P_{11}(0) = \mu I$ as in (54) and noting the fact that $K_{p,k} = F_k P_k H_k^*$ (Theorem 4), it is easy to see that as $k \to \infty$

$$
K_{p,k} \to \mu \begin{bmatrix} I \\ \sum_{j=0}^{k} A_s^j B_s h^*(k-j) \end{bmatrix}
$$
$$
\cdot \left(D_s h^*(k) + \sum_{j=0}^{k-1} C_s A_s^j B_s h^*(k-1-j) \right)^*
$$
$$
\to \mu \begin{bmatrix} I \\ \sum_{j=0}^{k} A_s^j B_s h^*(k-j) \end{bmatrix} h'(k) \quad (82)
$$

and therefore, the state estimate update in Theorem 4 becomes

$$
\begin{bmatrix} \hat{W}(k+1) \\ \hat{\Theta}(k+1) \end{bmatrix} = \begin{bmatrix} I & 0 \\ B_s h^*(k) & A_s \end{bmatrix} \begin{bmatrix} \hat{W}(k) \\ \hat{\Theta}(k) \end{bmatrix}
$$
$$
+ \mu \begin{bmatrix} h'(k) \\ \sum_{j=0}^{k} A_s^j B_s h^*(k-j)h'(k) \end{bmatrix}
$$
$$
\cdot \left(m(k) - D_s h^*(k)\hat{W}(k) - C_s \hat{\Theta}(k) \right). \quad (83)
$$

Thus, we have the following update for the weight vector

$$
\hat{W}(k+1) = \hat{W}(k) + \mu h'^*(k)
$$
$$
\cdot \left(m(k) - D_s h^*(k)\hat{W}(k) - C_s \hat{\Theta}(k) \right). \quad (84)
$$

Note that $m(k) = e(k) + y_{copy}(k)$ [see (66)], and hence, the difference between the limiting update rule of (84) (i.e., the prediction EBAF algorithm) and the classical FxLMS algorithm of (2) will be the error term used by these algorithms. More specifically, $e(k)$ in the FxLMS algorithm is replaced with the following *modified* error [using (65)]:

$$
e(k) + y_{copy}(k) - D_s h^*(k)\hat{W}(k) - C_s \hat{\Theta}(k)
$$
$$
= e(k) + C_s \Theta_{copy}(k) - C_s \hat{\Theta}(k). \quad (85)
$$

Note that if $y(k)$ is directly measurable, then $[e(k) + y(k) - D_s h^*(k)\hat{W}(k) - C_s \hat{\Theta}(k)]$ will be the modified error.

The condition for optimality of $\gamma = 1$ in the prediction case [see (56)] can also be simplified for stable LTI secondary path as $k \to \infty$. Rewriting the optimality condition for the prediction solution [see (56)] as

$$
\mu < \frac{1 - \alpha C_s A_s^k A_s^{*k} C_s^*}{h'^*(k)h'(k)} \quad (86)
$$

for a stable secondary path, $A_s^k \to 0$ as $k \to \infty$, and hence

$$
\mu < \frac{1}{h'^*(k)h'(k)} \text{ as } k \to \infty \quad (87)
$$

is the limiting condition for the optimality of $\gamma = 1$ in the prediction case. This is essentially a filtered version of the well-known LMS bound [22].

B. Filtering Solution and its Connection to the Normalized-FxLMS

In the filtered case, the gain vector is $K_{f,k} = F_k P_k H_k^*(I + H_k P_k H_k^*)$. In Section IX-A, we computed the limiting value for the quantity $F_k P_k H_k^*$ in (82). In a similar way, it can be shown that with $P_{11}(0) = \mu I$, as $K \to \infty$

$$
(I + H_k P_k H_k^*) \to (1 + \mu h'^*(k)h'(k)) \quad (88)
$$

and hence, the coefficient for the state estimate update in the filtering case becomes

$$
K_{f,k} \to \frac{\mu}{1 + \mu h'^*(k)h'(k)} \left(h'(k) \sum_{j=0}^{k} A_s^j B_s h^*(k-j)h'(k) \right)
$$
$$
\text{ as } k \to \infty. \quad (89)
$$
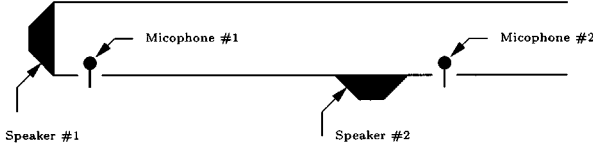
Fig. 5.   Schematic diagram of 1-D air duct.

Thus, the update rule for the weight vector in the filtering EBAF algorithm would be

$$\hat{W}(k+1) = \hat{W}(k) + \mu \frac{h'^*(k)}{(1 + \mu h'^*(k)h'(k))}$$

$$\cdot \left(m(k) - D_s h^*(k)\hat{W}(k) - C_s \hat{\Theta}(k) \right) \quad (90)$$

which is similar to the normalized-FxLMS algorithm [see (3)] in which the error signal is replaced with a modified error signal described by (85).

X. EXPERIMENTAL DATA AND SIMULATION RESULTS

This section examines the performance of the proposed EBAF algorithm for the active noise cancellation (ANC) problem in a one-dimensional (1-D) acoustic duct. The algorithm, however, is applicable for adaptive filter design in general, and the selection of the examples in this section are for pedagogical reasons. The application of the EBAF algorithm to a multichannel vibration isolation problem [15] and an adaptive equalization problem [24] are other instances of its successful application.

Fig. 5 shows the schematic diagram of the 1-D air duct that is used in our experiments. The control objective is to attenuate (cancel in the ideal case) the disturbance introduced into the duct by *Speaker #1* (primary noise source) at the position of *Microphone #2* (error sensor) by the control signal generated by *Speaker #2* (secondary source). Microphone #1 can be used to provide the reference signal for the adaptation algorithm. Clearly, Microphone #1 measurements are affected by both primary and secondary sources, and hence, if these measurements are used as the reference signal, the problem, which is commonly known as feedback contamination, has to be addressed.

We have used a dSPACE DS1102 DSP controller board (which includes TI's C31 DSP processor with 60 MHz clock rate and 128k of 32-bit RAM) and its Matlab 5 interface for real-time implementation of the algorithm. A state-space model (of order 10) is identified for this 1-D acoustic system. Note that of four identified transfer functions, only the transfer function from Speaker #2 to Microphone #2 (i.e., the secondary path) is required by the estimation-based adaptive algorithm. In this section, we will first provide experimental data that validate a corresponding simulation result. We then present more sophisticated simulations to study various aspects of the EBAF algorithm.

Fig. 6 shows the experimental data in a typical noise cancellation scenario, along with corresponding plots from a simulation that is designed to mimic that experiment. Here, we show the reading of Microphone#2 (i.e., the cancellation error) when an adaptive FIR filter of length 4 is used for noise cancellation. We have used a bandlimited white noise (noise power $= 0.008$) as measurement noise in our simulations. The sampling frequency
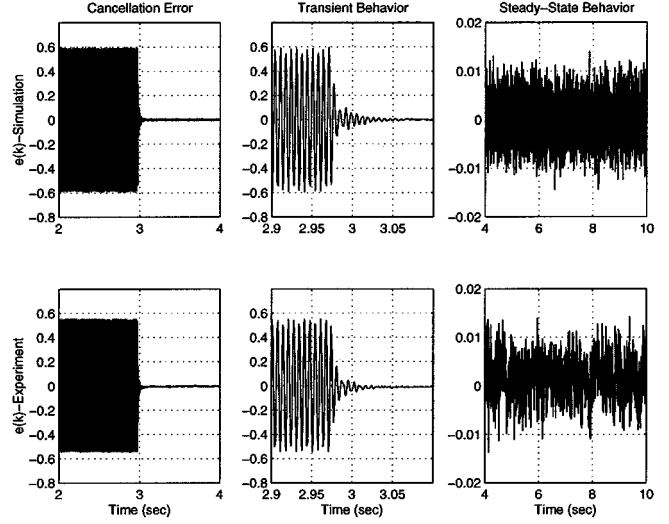

Fig. 6.   Validation of simulation results against experimental data for the noise cancellation problem with a single-tone primary disturbance at 150 Hz.
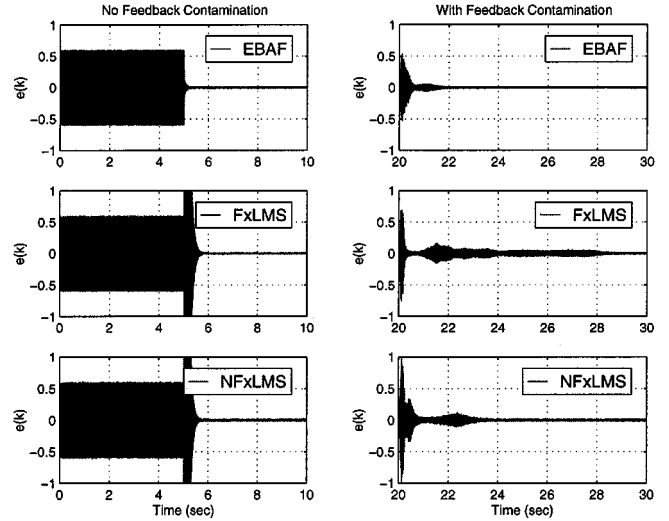

Fig. 7.   Performance comparison in the presence of feedback contamination when the primary source is a single tone at 150 Hz.

is 1000 Hz for both experiment and simulation. The primary source is a sinusoidal tone at 150 Hz, which is also available to the adaptation algorithm as the reference signal. We have selected $P_{11,0} = 0.05 I_{4 \times 4}$, $P_{12,0} = 0$, and $P_{22,0} = 0.005 I_{10 \times 10}$ to initialize the Riccati matrix in (14). Our experiment starts with adaptive controller off, and about 3 s later, the controller is turned on. The transient response of the adaptive FIR filter lasts for approximately 0.05 s. There is a reduction of 65 times in the magnitude of the error. The results from a corresponding Matlab simulation (with the same filter length and similar open loop error at 150 Hz) are also shown in Fig. 6. The transient behavior and the steady-state response in the simulation agree with the experimental data, thus assuring the validity of the setup for the simulations presented in this section.

In Fig. 7, we study the effect of feedback contamination (i.e., the contamination of the reference signal with the output of the adaptive FIR filter through some feedback path) when the primary source is a single tone. In [2], the subject of feedback contamination is discussed in detail, where relevant references to
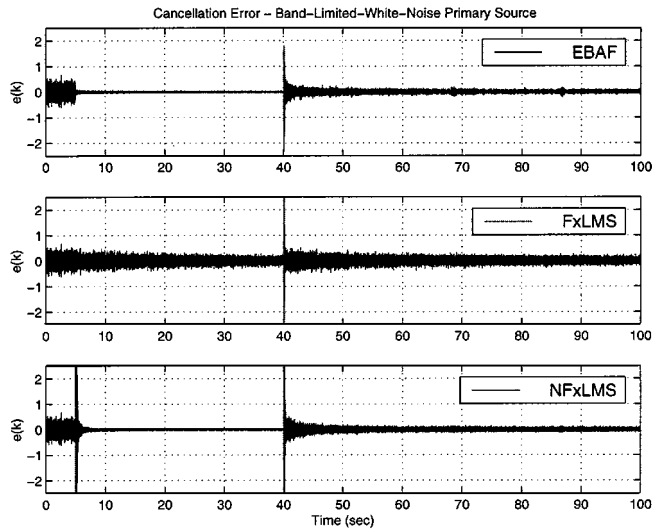
Fig. 8.　Performance comparison in the presence of feedback contamination when the primary source is a bandlimited white noise.
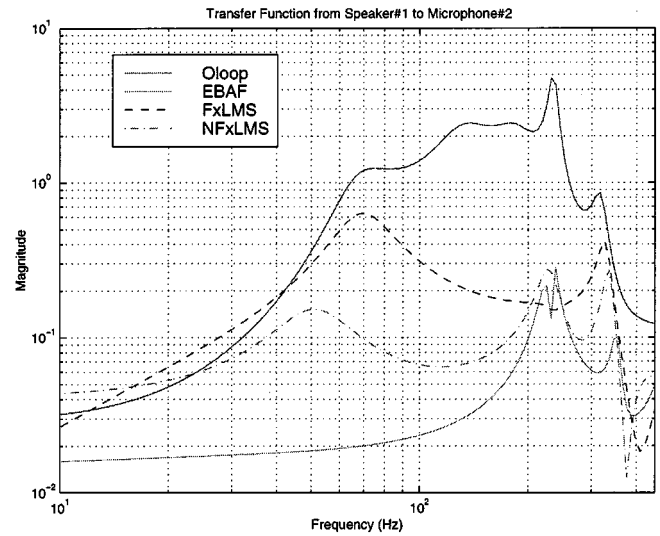

Fig. 9.　Closed-loop transfer function based on the steady-state performance of the EBAF and (N)FxLMS algorithms in the noise cancellation problem of Fig. 8.

the conventional solutions to this problem are also listed. Here, however, we do not intend to solve this problem. Instead, we will show that the proposed EBAF algorithm maintains superior performance [compared with FxLMS and normalized-FxLMS (NFxLMS) algorithms] when such a problem occurs and no additional information is furnished. It is worth mentioning that the EBAF algorithm provides a convenient framework in which the problem of feedback contamination can be addressed. This solution is discussed in [15]. Fig. 7 contains a typical response to feedback contamination for EBAF, FxLMS, and NFxLMS algorithms. For the first 5 s, the input to Speaker#2 is grounded [i.e., $u(k) = 0$ for $k \leq 5$]. Switching the controller on results in large transient behavior in the case of FxLMS and NFxLMS, whereas for the EBAF algorithm, the transient behavior does not display the undesirable overshoot. We have tested many different operation scenarios (with various filter lengths, and adaptation rates), and this observation holds true in all cases tested. For the next 15 s, the primary source is directly available to all adaptive algorithms, and the steady-state performance is virtually the same. From $k = 20$ on, we use the output of Microphone#1 (which is contaminated by the output of the FIR filter) as the reference signal. Once again, Fig. 7 shows a typical observation. Note that in the case of FxLMS and NFxLMS, the adaptation rate must be kept small enough to avoid unstable behavior when the switch to contaminated reference signal takes place. The superior robustness of the EBAF algorithm allows for a faster convergence in the face of this feedback contamination. For the results in Fig. 7, the length of the adaptive FIR filter (for all three algorithms) is 24. For the EBAF algorithm, $P_{11,0} = 0.005 I_{24 \times 24}$, $P_{22,0} = 0.0005 I_{10 \times 10}$, and $P_{12,0} = 0$. For FxLMS and NFxLMS algorithms, the adaptation rates are 0.005 and 0.025, respectively.

Fig. 8 considers the effect of feedback contamination in a wide band noise cancellation process. As in Fig. 7, the controller is off for the first 5 s. For $5 \leq t \leq 40$ s, the controller has full access to the reference signal, whereas at $t = 40$ s, the reference signal is switched to the measurements of the Microphone#1 (hence, contaminated with the feedback from Speaker#2). For the results in

Fig. 8, the length of the adaptive FIR filter is 32. For the EBAF algorithm, $P_{11,0} = 0.05 I_{32 \times =32}$, $P_{22,0} = 0.005 I_{10 \times 10}$, and $P_{12,0} = 0$. For FxLMS and NFxLMS algorithms, the adaptation rates are 0.0005 and 0.1, respectively. The FxLMS algorithm becomes unstable for faster adaptation rates, hence forcing slow convergence (i.e., lower control bandwidth). For NFxLMS, the normalization of the adaptation rate by the norm of the reference vector (a vector of length 32 in this case) prevents unstable behavior. The response of the algorithm under feedback contamination is, however, still slower than EBAF algorithm. Furthermore, the oscillations in the cancellation error due to the switching between modes of operation are significantly higher when compared twith the oscillations in the EBAF case.

Fig. 9 shows a closed-loop performance comparison for wideband noise cancellation. The EBAF algorithm outperforms FxLMS and normalized-FxLMS adaptive algorithms, even though the same level of information is made available to all three adaptation schemes. For the result presented here, the length of the FIR filter (for all three approaches) is 32, and the bandlimited white noise that is used as the primary source is available as the reference signal. Since the frequency response is calculated based on the steady-state data, the adaptation rate of the algorithms is not relevant. Measurement noise for all three simulations is a bandlimited white noise with power 0.008, as before.

XI. Conclusion

We have approached the adaptive control problem from an estimation point of view. More specifically, we have shown that for a common formulation of the adaptive control problem, an equivalent estimation interpretation exists. We have then constructed a standard H_∞ estimation problem that corresponds to the original adaptive control problem and have justified our choice of estimation criterion. We have also derived the H_∞-optimal filtering/prediction solutions and proved that the optimal energy gain is unity. We have then simplified the filtering/prediction solutions and explained how these solutions form the

foundation for an estimation-based adaptive filtering (EBAF) algorithm. Meanwhile, we have justified why the real-time implementation of EBAF algorithm is feasible.

We have outlined an implementation scheme for the new algorithm and have derived a corresponding performance bound. We have shown that the classical FxLMS (normalized-FxLMS) adaptive algorithms are approximations to the limiting behavior of the proposed EBAF algorithm. We have shown that the EBAF algorithm displays improved performance when compared with commonly used FxLMS and normalized-FxLMS algorithms. We have also verified our simulations by conducting a noise cancellation experiment and showing that the experimental data reasonably match a corresponding simulation.

The systematic nature of the proposed EBAF algorithm can serve as the first step toward methodical optimization of now predetermined parameters of the FIR filter (such as filter length or adaptation rate). Furthermore, the analysis of the various aspects of the algorithm directly benefits from the advances in robust estimation theory. Finally, more efficient implementation schemes can further reduce computational complexity of the algorithm.

REFERENCES

[1] B. Widrow and S. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[2] S. Kuo and D. Morgan, *Adaptive Noise Control Systems*. New York: Wiley, 1996.

[3] T. Shan and T. Kailath, "Adaptive algorithms with an automatic gain control feature," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 122–127, Jan. 1988.

[4] R. Gitlin, H. Meadors, and S. Weinstein, "The tap-leakage algorithm: An algorithm for the stable operation of a digitally implemented, fractional adaptive space equalizer," *Bell Syst. Tech. J.*, vol. 61, pp. 1817–1839, Oct. 1982.

[5] R. Harris, D. Chabries, and F. Bishop, "A variable step adaptive filter algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 309–316, Apr. 1986.

[6] D. Morgan, "An analysis of multiple correlation cancellation loops with a filter in the auxiliary path," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 454–467, Aug. 1980.

[7] K. Astrom and B. Wittenmark, *Adaptive Control*. Reading, MA: Addison-Wesley, 1989.

[8] B. Widrow, "Adaptive filters I: Fundamentals," Electron. Labs., Stanford Univ., Stanford, CA, 1966.

[9] J. Glover, "Adaptive noise cancellation applied to sinusoidal interferences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, pp. 484–491, Dec. 1977.

[10] B. W. M. Dentino and J. McCool, "Adaptive filtering in frequency domain," *Proc. IEEE*, vol. 66, pp. 1658–1659, Dec. 1978.

[11] S. Elliott, I. Stothers, and P. Nelson, "A multiple error LMS algorithm and its application to the active control of sound and vibration," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1423–1434, Oct. 1987.

[12] N. Bershad and O. Macchi, "Adaptive recovery of a chirped sinusoid in noise— Part 2: Performance of the LMS algorithm," *IEEE Trans. Signal Processing*, vol. 39, pp. 595–602, Mar. 1991.

[13] P. Feintuch, N. Bershad, and A. Lo, "A frequency domain model for filtered LMS algorithm-stability analysis, design, and elimination of the training mode," *IEEE Trans. Signal Processing*, vol. 41, pp. 1518–1531, Apr. 1993.

[14] E. Bjarnason, "Analysis of the filtered-X LMS algorithm," *IEEE Trans. Speech Audio Processing*, vol. 3, pp. 504–514, Nov. 1995.

[15] B. Sayyar-Rodsari, "Estimation-based adaptive filtering and control," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1999.

[16] K. Narendra and A. Annaswamy, *Stable Adaptive Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[17] G. Zames, "Feedback optimal sensitivity: Model preference transformation, multiplicative seminorms and approximate inverses," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 301–320, 1981.

[18] J. Doyle, K. Glover, P. Khargonekar, and B. Francis, "State-space solutions to standard H_2 and H_∞ control problems," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 831–847, Aug. 1989.

[19] M. Green and D. Limebeer, *Linear Robust Control*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[20] P. Khargonekar and K. Nagpal, "Filtering and smoothing in an H_∞ setting," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 151–166, 1991.

[21] Y. Shaked and Y. Theodor, "H_∞-optimal estimation: A tutorial," in *Proc. IEEE Conf. Decision Contr.*, vol. 2, Dec. 1992.

[22] B. Hassibi, A. Sayed, and T. Kailath, "H_∞ optimality of the LMS algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 267–280, Feb. 1996.

[23] B. Hassibi, A. Sayed, and T. Kailath, "Indefinite quadratic estimation and control," in *SIAM Studies Appl. Math.*, 1998.

[24] A. Tehrani, B. Sayyarrodsari, B. Hassibi, J. How, and J. Cioffi, "Estimation-based synthesis of H_∞-optimal adaptive equalizers over wireless channels," in *Proc. Globecom*, vol. 1a, 1999, pp. 457–461.

Bijan Sayyarrodsari (M'00) received the B.Sc. degree from Sharif University of Technology, Tehran, Iran, in 1990. He received the M.Sc. and Ph.D. degrees from Stanford University, Stanford, CA, in 1996 and 1999, respectively.

He is now a Research Scientist with Pavilion Technologies, Inc., Austin, TX. His research interests include robust adaptive control and estimation and hybrid system control and modeling with the emphasis on their application in process industry and financial systems.

Dr. Sayyarrodsari is a Member of SIAM.

Jonathan P. How received the B.A.Sc. degree in engineering science (aerospace option) from the University of Toronto, Toronto, ON, Canada, in 1987 and the S.M. and Ph.D. degrees in aeronautics and astronautics from the Massachusetts Institute of Technology (MIT), Cambridge, in 1990 and 1993, respectively.

He is currently an Associate Professor with the Department of Aeronautics and Astronautics, MIT. He studied for two years at MIT as a Postdoctoral Associate in charge of the design and analysis of robust controllers for the Middeck Active Control Experiment (MACE), which flew on-board the Space Shuttle Endeavour in March 1995. Prior to joining MIT in 2000, he worked for five years as an Assistant Professor with the Department of Aeronautics and Astronautics, Stanford University, Stanford, CA. His current research focuses on various aspects of spacecraft navigation and control, including GPS sensing for formation-flying vehicles, theoretical analysis and synthesis of robust, hybrid, and adaptive controllers, and modeling and experimental control of vibrations in flexible systems.

Dr. How is a Senior Member of AIAA and is active in the IEEE and ION.

Babak Hassibi was born in Tehran, Iran, in 1967. He received the B.S. degree from the University of Tehran in 1989 and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1993 and 1996, respectively, all in electrical engineering.

From October 1996 to October 1998, he was a Research Associate with the Information Systems Laboratory, Stanford University, and since November 1998, he has been Member of Technical Staff at Bell Laboratories, Murray Hill, NJ. Starting January 2001, he will be an Assistant Professor of electrical engineering at the California Institute of Technology, Pasadena. He has also held short-term appointments at Ricoh California Research Center, the Indian Institute of Science, and Linköping University, Linköping, Sweden. His research interests include wireless communications, robust estimation and control, adaptive signal processing and linear algebra. He is the coauthor of the books *Indefinite Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories* (New York: SIAM, 1999) and *Linear Estimation* (Englewood Cliffs, NJ: Prentice-Hall, 2000). He was also the recipient of the 1999 O. Hugo Schuck best paper award of the American Automatic Control Council.

Alain Carrier received the Ph.D. degree in aeronautics and astronautics from Stanford University, Stanford, CA, in 1990.

Since then, he has been working for the Lockheed Martin Advanced Technology Center, Palo Alto, CA, leading applied research and optical-precision instrumentation design, modeling, and control. He lead the development of several actively controlled electromechanical systems from concept to hardware demonstration, including actively controlled segmented optics, secondary and fast steering mirrors for astronomical telescopes, zero-G slew suspensions for space structures, active and passive vibration isolators, smart actuators, and a latch mechanism actuated by Shape Memory Alloy springs for which he owns a patent. He is the author of the Principal Gain Tracking (PGT), which is a novel testing and system identification technique for high-modal-density lightly damped structures. He currently leads the development of the pointing control system for HIRDLS (an earth observing radiometer) and the development and experimental demonstration of adaptive control techniques for vibration isolation. His research interests are in isolation, control, and passive damping of broadband and periodic mechanical vibrations for optical instruments; subarc-second optical pointing and beam control for earth observing, laser communication, and astronomical instruments; actuators and sensors for structural control; dynamics modeling of space structures and instruments; attitude control, stationkeeping, slews and orbital maneuvers of spacecrafts, and "sciencecrafts."