

Correspondence

Analysis of an LMS Algorithm for Unbiased Impulse Response Estimation

H. C. So and Y. T. Chan

Abstract—In this correspondence, a least mean squares (LMS)-based algorithm is devised for unbiased system identification in the presence of white input and output noise, assuming that the ratio of the noise powers is known. The proposed approach aims to minimize the mean square value of the equation-error function under a constant-norm constraint and is equivalent to minimizing a modified mean square error (MSE) function. An analysis of the algorithm shows that the estimates will converge to the true values in the mean sense. The variances of the parameter estimates are also available. Computer simulations are included to corroborate the theoretical development and to evaluate the impulse response estimation performance of the LMS algorithm under different conditions.

Index Terms—Adaptive filter, bias removal, system identification.

I. INTRODUCTION

Estimation of the impulse response of an unknown system from its measured input and output has found many applications in signal processing, communications, and control [1]–[3]. Many estimation algorithms based on least squares (LS) and least mean squares (LMS) techniques [4]–[9] have been developed for accurate system parameter estimation if additive noise exists only at the output. However, noise-free input measurements are difficult to obtain in many practical situations because these signals are also corrupted by quantization noise, self-noise, and/or other additive noise [10], and neglecting the presence of input noise will lead to biased impulse response estimates.

The total least squares (TLS) approach [11] seems to be an appropriate choice for system identification with noisy input and output because it considers both stimulation and response errors. Computationally attractive adaptive impulse response estimation schemes employing the TLS criterion include the recursive total least squares (RTLTS) [10], constrained anti-Hebbian (CAH) [12], and total least mean squares (TLMS) [13] algorithms. Nevertheless, these methods are restricted for parameter estimation of finite impulse response (FIR) systems and assume that the input and output noise powers are identical. For the general case of infinite impulse response (IIR) system modeling, the Koopmans–Levin method [14], which is based on spectral decomposition of a covariance matrix formed using the input-output data and a modified LS algorithm [15] that requires estimation of the measurement noise variances, can be used to provide reliable estimates but they involve intensive computations. In this correspondence, we develop an LMS approach for unbiased impulse response estimation for which real-time implementation is allowed, assuming that the power ratio of the input and output interference is

Manuscript received May 2, 2002; revised January 29, 2003. The associate editor coordinating the review of this paper and approving it for publication was Prof. Fredrik Gustafsson.

H. C. So is with the Department of Computer Engineering and Information Technology, City University of Hong Kong, Kowloon, Hong Kong (e-mail: itheso@cityu.edu.hk).

Y. T. Chan is with the Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong (e-mail: ytchan@ee.cuhk.edu.hk).

Digital Object Identifier 10.1109/TSP.2003.812747

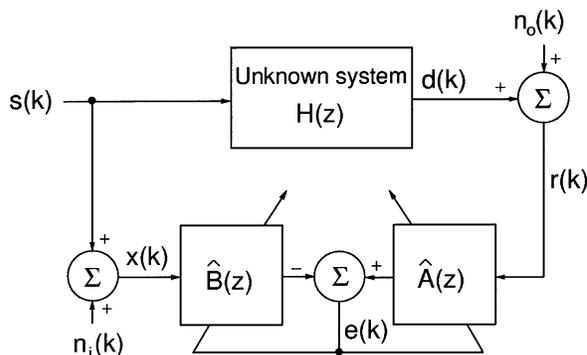


Fig. 1. System block diagram for equation-error IIR filtering.

known [14]. In estimating IIR system parameters, the equation-error configuration is employed because it has advantages of unimodal error surface and guaranteed system stability over the output-error approach, although direct minimization of the mean square value of the equation-error cannot provide unbiased parameter estimates in input and/or output noise [7]–[9].

The organization of the paper is as follows. In Section II, the problem of impulse response estimation using noisy input and output measurements is formulated. It is shown that bias-free estimation is possible if the minimization is subject to a constant-norm constraint. The constrained problem is then converted into an unconstrained minimization of a weighted mean square error (MSE) function by an LMS algorithm. Learning characteristics and steady-state MSEs of the impulse response estimates are derived in Section III. Simulation results are presented in Section IV to corroborate the theoretical analyses and to evaluate the system identification performance of the new approach under different conditions. Finally, conclusions are drawn in Section V.

II. BIAS-FREE LMS ALGORITHM

The block diagram for adaptive equation-error IIR filtering is in Fig. 1. Without loss of generality, we consider that the unknown system $H(z)$ is stable as well as causal and has the form

$$H(z) = \frac{B(z)}{A(z)} \tag{1}$$

where

$$A(z) = 1 + \sum_{i=1}^{M-1} a_i z^{-i}$$

$$B(z) = \sum_{j=0}^{N-1} b_j z^{-j}. \tag{2}$$

The orders of the denominator and numerator polynomials M and N , respectively, are assumed to be known. The observed system input and output are

$$x(k) = s(k) + n_i(k)$$

$$r(k) = d(k) + n_o(k) \tag{3}$$

where $s(k)$ and $d(k)$ denote the noise-free input and output, respectively, whereas $n_i(k)$ and $n_o(k)$ represent the input and output measurement noise that are independent of $s(k)$. It is assumed that $n_i(k)$

and $n_o(k)$ are uncorrelated white processes with variances $\sigma_{n_i}^2$ and $\sigma_{n_o}^2$, respectively, and the ratio of the noise powers, say $\gamma = \sigma_{n_i}^2 / \sigma_{n_o}^2$, is available. Our goal is to find a_i , $i = 1, 2, \dots, M-1$ and b_j , $j = 0, 1, \dots, N-1$ from $x(k)$ and $r(k)$ using the adaptive FIR filters $\hat{A}(z) = 1 + \sum_{i=1}^{M-1} \hat{a}_i z^{-i}$ and $\hat{B}(z) = \sum_{j=0}^{N-1} \hat{b}_j z^{-j}$, where $\{\hat{a}_i\}$ and $\{\hat{b}_j\}$ are the parameter estimates.

In the equation-error formulation, the output error function $e(k)$ is computed from

$$e(k) = r(k) + \sum_{i=1}^{M-1} \hat{a}_i r(k-i) - \sum_{j=0}^{N-1} \hat{b}_j x(k-j). \quad (4)$$

It is easy to show that the mean square value of $e(k)$, $E\{e^2(k)\}$ is

$$E\{e^2(k)\} = E \left\{ \left(d(k) + \sum_{i=1}^{M-1} \hat{a}_i d(k-i) - \sum_{j=0}^{N-1} \hat{b}_j s(k-j) \right)^2 \right\} + \sigma_{n_o}^2 \left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2 \right). \quad (5)$$

In the presence of input and/or output noise, it is obvious that minimizing $E\{e^2(k)\}$ with respect to $\{\hat{a}_i\}$ and $\{\hat{b}_j\}$ cannot give unbiased impulse response estimates because (5) contains the noise components. Using the idea of [6], if we minimize $E\{e^2(k)\}$ subject to the constraint that $\left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2\right)$ is a positive constant, say G , then the desired solution can be attained. This optimization problem is equivalent to

minimize

$$\frac{E\{e^2(k)\}}{G} \quad \text{or} \quad \frac{1}{G} E \left\{ \left(d(k) + \sum_{i=1}^{M-1} \hat{a}_i d(k-i) - \sum_{j=0}^{N-1} \hat{b}_j s(k-j) \right)^2 \right\} + \frac{\sigma_{n_o}^2}{G} \left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2 \right)$$

subject to

$$\frac{1}{G} \left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2 \right) = 1. \quad (6)$$

Since the variables $\{a_i\}$ and $\{b_j\}$ in (6) appear in the first and second degrees only, the error surface is quadratic, and thus, it is unimodal [5]. This implies that gradient search approach can be utilized for the constrained minimization. From (6)

$$G = 1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2. \quad (7)$$

As a result, (6) is identical to an unconstrained optimization of minimizing a modified error function $E\{\zeta^2(k)\}$ of the form

$$E\{\zeta^2(k)\} = \frac{E\{e^2(k)\}}{1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2}. \quad (8)$$

In our study, the computationally attractive LMS algorithm is used to estimate $H(z)$ on a sample-by-sample basis. Based on (8), the instantaneous value of $\zeta(k)$ is

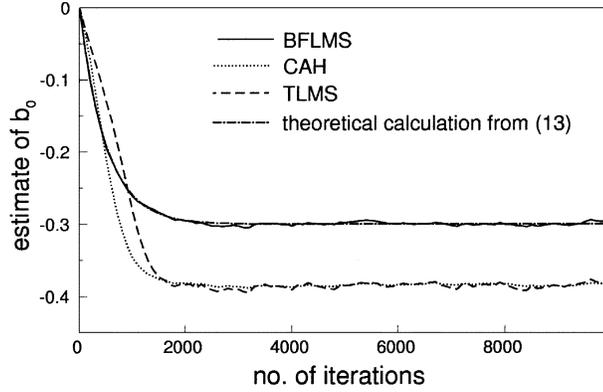
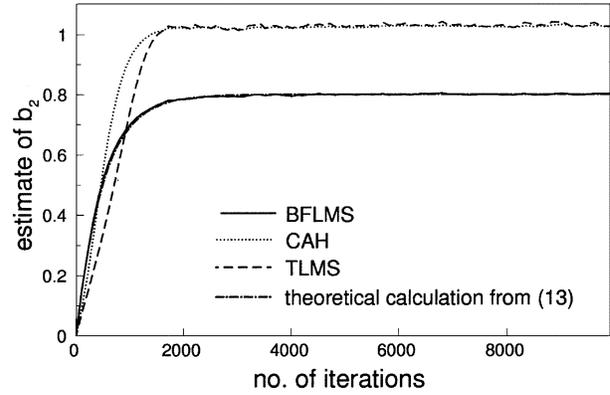
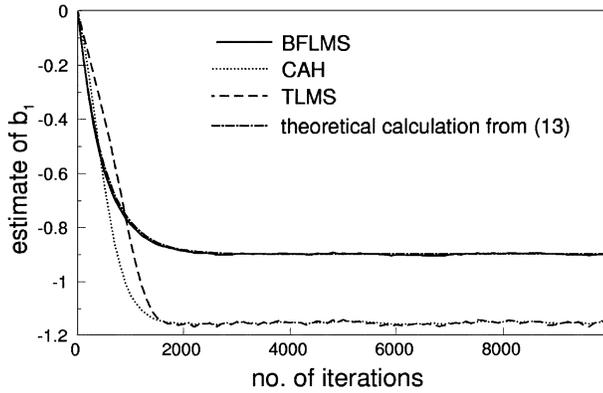
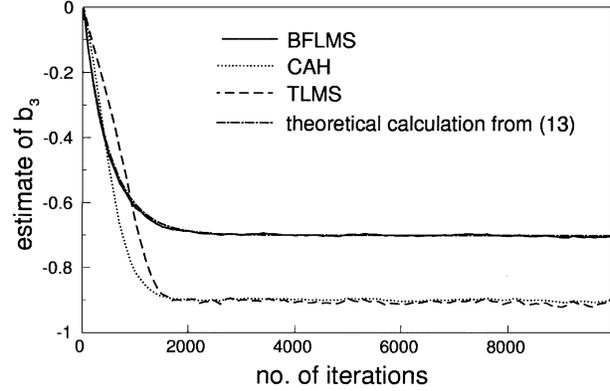
$$\zeta(k) = \frac{e(k)}{\sqrt{1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k)}} \quad (9)$$

where $\hat{a}_i(k)$, $i = 1, 2, \dots, M-1$, $\hat{b}_j(k)$, and $j = 0, 1, \dots, N-1$ denote the estimates of $\{a_i\}$ and $\{b_j\}$ at time k , whereas $e(k)$ is calculated using the instantaneous parameter estimates. The bias-free least mean squares (BFLMS) updating equations for unbiased system identification are then derived as in (10) and (11), shown at the bottom of the page, where μ_a and μ_b are positive scalars that govern convergence rate and stability of the adaptive algorithm. We note that (10) and (11) are a modification of the equation-error scheme where bias-removal terms are added to the update of $\{\hat{a}_i(k)\}$ and $\{\hat{b}_j(k)\}$. At each sampling interval, only $(4M + 4N - 2)$ multiplications, one division, and $(4M + 4N - 4)$ additions are required in the BFLMS algorithm.

$$\begin{aligned} \hat{a}_i(k+1) &= \hat{a}_i(k) - \frac{\mu_a \left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k) \right)}{2} \frac{\partial \zeta^2(k)}{\partial \hat{a}_i(k)} \\ &= \hat{a}_i(k) - \mu_a e(k) \left[r(k-i) - \hat{a}_i(k) \frac{e(k)}{1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k)} \right] \\ & \quad i = 1, 2, \dots, M-1 \end{aligned} \quad (10)$$

and

$$\begin{aligned} \hat{b}_j(k+1) &= \hat{b}_j(k) - \frac{\mu_b \left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k) \right)}{2} \frac{\partial \zeta^2(k)}{\partial \hat{b}_j(k)} \\ &= \hat{b}_j(k) + \mu_b e(k) \left[x(k-j) + \gamma \hat{b}_j(k) \frac{e(k)}{1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k)} \right] \\ & \quad j = 0, 1, \dots, N-1 \end{aligned} \quad (11)$$

Fig. 2. Estimates of b_0 using different methods in FIR filtering.Fig. 4. Estimates of b_2 using different methods in FIR filtering.Fig. 3. Estimates of b_1 using different methods in FIR filtering.Fig. 5. Estimates of b_3 using different methods in FIR filtering.

III. PERFORMANCE ANALYSIS OF BFLMS ALGORITHM

Assume the fluctuations in the coefficient errors are small so that the elements of $\{\hat{a}_i(k)\}$ and $\{\hat{b}_j(k)\}$ can be replaced by their true values and that $E\{e^2(k)\} \approx e^2(k)$. Then, $e^2(k)/(1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k)) \approx \sigma_{n_o}^2$. Let

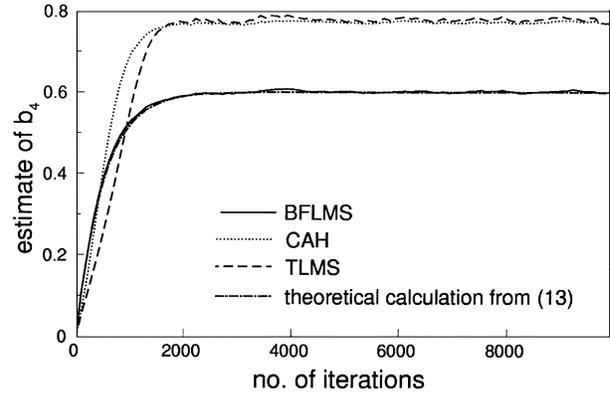
$$\begin{aligned} \alpha(k) &= [E\{\hat{a}_1(k)\} - a_1, E\{\hat{a}_2(k)\} - a_2, \dots \\ &\quad E\{\hat{a}_{M-1}(k)\} - a_{M-1}]^T \\ \beta(k) &= [E\{\hat{b}_0(k)\} - b_0, E\{\hat{b}_1(k)\} - b_1, \dots \\ &\quad E\{\hat{b}_{N-1}(k)\} - b_{N-1}]^T \\ \mathbf{d}(k) &= [d(k-1), d(k-2), \dots, d(k-M+1)]^T \\ \mathbf{s}(k) &= [s(k), s(k-1), \dots, s(k-N+1)]^T \end{aligned} \quad (12)$$

where T denotes the transpose operation, as well as $\mathbf{R}_{dd} = E\{\mathbf{d}(k)\mathbf{d}^T(k)\}$, $\mathbf{R}_{ds} = E\{\mathbf{d}(k)\mathbf{s}^T(k)\}$, $\mathbf{R}_{ss} = E\{\mathbf{s}(k)\mathbf{s}^T(k)\}$, and $\kappa = \mu_b/\mu_a$. Together with the independence assumption [5] and taking expectations on both sides of (10) and (11), the learning behavior of the BFLMS estimates can be obtained from

$$\begin{bmatrix} \alpha(k+1) \\ \beta(k+1) \end{bmatrix} = [\mathbf{I} - \mu_a \mathbf{R}] \begin{bmatrix} \alpha(k) \\ \beta(k) \end{bmatrix} \quad (13)$$

where

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{dd} & \mathbf{R}_{ds} \\ \kappa \mathbf{R}_{ds}^T & \kappa \mathbf{R}_{ss} \end{bmatrix}. \quad (14)$$

Fig. 6. Estimates of b_4 using different methods in FIR filtering.

and \mathbf{I} is the $(M+N-1) \times (M+N-1)$ identity matrix. From (13), it is observed that closed-form solutions for $E\{\hat{a}_i(k)\}$ and $E\{\hat{b}_j(k)\}$ are not generally available because the convergence characteristic of each parameter estimate is dependent on the auto-correlation and cross-correlation functions, although they can be determined by brute force if the signal statistics are known. Nevertheless, we can see that $\hat{a}_i(k) \rightarrow a_i$, $i = 1, 2, \dots, M-1$, and $\hat{b}_j(k) \rightarrow b_j$, $j = 1, 2, \dots, N-1$, as $k \rightarrow \infty$ in the mean sense, provided that $0 < \mu_a < 2/\lambda_{\max}$ and $0 < \mu_b < 2\kappa/\lambda_{\max}$, where λ_{\max} is the largest eigenvalue of \mathbf{R} , that is, there is convergence.

By squaring (10) and (11), taking the expected values on both sides, and then considering the steady-state with sufficiently small μ_a and μ_b , the variances of the impulse response estimates, which are denoted by $\{\text{var}(\hat{a}_i)\}$ and $\{\text{var}(\hat{b}_j)\}$, can be approximated as in (15) and (16),

TABLE I
MEAN IMPULSE RESPONSE ESTIMATES OF DIFFERENT METHODS IN FIR FILTERING

Filter Coefficient		b_0	b_1	b_2	b_3	b_4
Actual Value		-0.3	-0.9	0.8	-0.7	0.6
$\sigma_{n_i}^2 = 0.1$	BFLMS	-2.994×10^{-1}	-9.009×10^{-1}	8.000×10^{-1}	-7.004×10^{-1}	6.008×10^{-1}
	CAH	-3.837×10^{-1}	-1.154×10^0	1.026×10^0	-8.993×10^{-1}	7.721×10^{-1}
$\sigma_{n_o}^2 = 1.0$	TLMS	-3.844×10^{-1}	-1.156×10^0	1.031×10^0	-9.071×10^{-1}	7.784×10^{-1}
$\sigma_{n_i}^2 = 0.1$	BFLMS	-2.996×10^{-1}	-9.008×10^{-1}	8.002×10^{-1}	-7.007×10^{-1}	6.002×10^{-1}
	CAH	-2.995×10^{-1}	-9.001×10^{-1}	7.999×10^{-1}	-7.003×10^{-1}	6.002×10^{-1}
$\sigma_{n_o}^2 = 0.1$	TLMS	-2.996×10^{-1}	-9.024×10^{-1}	8.028×10^{-1}	-7.051×10^{-1}	6.036×10^{-1}
$\sigma_{n_i}^2 = 1.0$	BFLMS	-3.042×10^{-1}	-9.087×10^{-1}	8.047×10^{-1}	-7.094×10^{-1}	6.059×10^{-1}
	CAH	-2.979×10^{-1}	-8.983×10^{-1}	7.995×10^{-1}	-7.033×10^{-1}	6.020×10^{-1}
$\sigma_{n_o}^2 = 1.0$	TLMS	-3.001×10^{-1}	-9.007×10^{-1}	8.024×10^{-1}	-7.065×10^{-1}	6.053×10^{-1}
$\sigma_{n_i}^2 = 1.0$	BFLMS	-3.050×10^{-1}	-9.077×10^{-1}	8.044×10^{-1}	-7.089×10^{-1}	6.042×10^{-1}
	CAH	-2.263×10^{-1}	-6.822×10^{-1}	6.066×10^{-1}	-5.325×10^{-1}	4.547×10^{-1}
$\sigma_{n_o}^2 = 0.1$	TLMS	-2.279×10^{-1}	-6.839×10^{-1}	6.080×10^{-1}	-5.345×10^{-1}	4.563×10^{-1}

TABLE II
MSES OF THE PARAMETER ESTIMATES FOR DIFFERENT METHODS IN FIR FILTERING

Filter Coefficient		b_0	b_1	b_2	b_3	b_4
$\sigma_{n_i}^2 = 0.1$	BFLMS	1.435×10^{-3}	1.385×10^{-3}	1.314×10^{-3}	1.321×10^{-3}	1.452×10^{-3}
	BFLMS (pred.)	1.362×10^{-3}	1.355×10^{-3}	1.357×10^{-3}	1.358×10^{-3}	1.359×10^{-3}
$\sigma_{n_o}^2 = 1.0$	CAH	8.296×10^{-3}	6.689×10^{-2}	5.330×10^{-2}	4.175×10^{-2}	3.144×10^{-2}
	TLMS	1.107×10^{-2}	7.303×10^{-2}	6.029×10^{-2}	4.935×10^{-2}	3.742×10^{-2}
$\sigma_{n_i}^2 = 0.1$	BFLMS	3.950×10^{-4}	3.708×10^{-4}	3.532×10^{-4}	3.502×10^{-4}	3.833×10^{-4}
	BFLMS (pred.)	3.720×10^{-4}	3.648×10^{-4}	3.665×10^{-4}	3.680×10^{-4}	3.693×10^{-4}
$\sigma_{n_o}^2 = 0.1$	CAH	2.085×10^{-4}	4.161×10^{-4}	3.771×10^{-4}	3.340×10^{-4}	3.127×10^{-4}
	TLMS	2.091×10^{-3}	4.435×10^{-3}	4.091×10^{-3}	3.709×10^{-3}	3.189×10^{-3}
$\sigma_{n_i}^2 = 1.0$	BFLMS	7.741×10^{-3}	5.051×10^{-3}	5.221×10^{-3}	5.398×10^{-3}	5.673×10^{-3}
	BFLMS (pred.)	6.690×10^{-3}	5.970×10^{-3}	6.140×10^{-3}	6.290×10^{-3}	6.420×10^{-3}
$\sigma_{n_o}^2 = 1.0$	CAH	3.803×10^{-3}	5.059×10^{-3}	4.920×10^{-3}	4.520×10^{-3}	4.228×10^{-3}
	TLMS	3.813×10^{-3}	5.139×10^{-3}	4.997×10^{-3}	4.627×10^{-3}	4.295×10^{-3}
$\sigma_{n_i}^2 = 1.0$	BFLMS	6.053×10^{-3}	3.371×10^{-3}	3.305×10^{-3}	3.397×10^{-3}	3.465×10^{-3}
	BFLMS (pred.)	4.890×10^{-3}	4.170×10^{-3}	4.340×10^{-3}	4.490×10^{-3}	4.620×10^{-3}
$\sigma_{n_o}^2 = 0.1$	CAH	6.699×10^{-3}	4.929×10^{-2}	3.914×10^{-2}	2.967×10^{-2}	2.260×10^{-2}
	TLMS	7.099×10^{-3}	4.950×10^{-2}	3.954×10^{-2}	2.986×10^{-2}	2.288×10^{-2}

IV. NUMERICAL EXAMPLES

shown at the bottom of the next page. We observe that the MSEs of $\{\hat{a}_i(k)\}$ and $\{\hat{b}_j(k)\}$ are characterized by the step sizes, signal and noise powers as well as the system parameters of $H(z)$.

This section contains simulation results to corroborate the theoretical derivations and to evaluate the performance of the BFLMS al-

gorithm for estimating FIR and IIR system parameters using noisy measurements. For FIR system identification, performance comparison was also made with two LMS-style TLS algorithms, namely, the CAH and TLMS methods. The noise-free input $s(k)$ was a white Gaussian process of unity power, and the measurement noises $n_i(k)$ and $n_o(k)$ were independent white uniform random variables. All simulation results provided were averages of 200 independent runs.

In the first test, the system to be estimated was an FIR filter with $N = 5$ characterized by $b_0 = -0.3, b_1 = -0.9, b_2 = 0.8, b_3 = -0.7,$ and $b_4 = 0.6$. In this case, $\hat{A}(z) = 1$, and only (11) was involved in the proposed algorithm. All adaptive filter coefficients were set to zero initially, except that the initial weight for $r(k)$ in the TLMS method was -1 . We assigned $\mu_a = \mu_b = 0.002$, and the step sizes for the other two algorithms were chosen to be 0.001 so that all schemes had similar convergence times. Figs. 2–6 show the learning behavior of the impulse response estimates for different methods when $\sigma_{n_i}^2 = 0.1$ and $\sigma_{n_o}^2 = 1.0$. We see that all parameter estimates converged approximately at the 2000th iteration, and the unbiasedness of the BFLMS algorithm was demonstrated. The convergence characteristics of the BFLMS estimates also agreed with the predicted trajectories based on (13). On the other hand, the CAH and TLMS could not give unbiased estimation, and interestingly, they provided almost identical steady-state filter coefficient estimates. This test was repeated with three more noise conditions, namely, $\{\sigma_{n_i}^2 = 0.1, \sigma_{n_o}^2 = 0.1\}$, $\{\sigma_{n_i}^2 = 1.0, \sigma_{n_o}^2 = 1.0\}$, and $\{\sigma_{n_i}^2 = 1.0, \sigma_{n_o}^2 = 0.1\}$, and Tables I and II tabulate the means and MSEs of the impulse response estimates, respectively, for the three approaches in all four cases. From Table I, it can be observed that the BFLMS method always attained unbiased system parameter estimation, whereas the TLS-based algorithms gave accurate estimates only when $\sigma_{n_i}^2 = \sigma_{n_o}^2$. Furthermore, we noticed that the CAH and TLMS estimates were also very close for $\sigma_{n_i}^2 = 1.0$ and $\sigma_{n_o}^2 = 0.1$, and their biases varied with the noise powers. From Table II, we see that the variances of the BFLMS estimates conformed with the theoretical calculation based on (16) for all noise conditions, and it achieved minimum MSE estimation for the cases when $\sigma_{n_i}^2 \neq \sigma_{n_o}^2$ because the biases of the CAH and TLMS estimates had contributed to their MSEs. For $\sigma_{n_i}^2 = \sigma_{n_o}^2 = 1.0$, all three methods provided similar values of MSEs, whereas for $\sigma_{n_i}^2 = \sigma_{n_o}^2 = 0.1$, the measured MSEs obtained from the BFLMS and CAH algorithms were also comparable and were far fewer than those of the TLMS scheme. As expected, the MSEs of all three techniques increased with the values of the noise variances.

In the second test, we evaluated the system identification performance of the BFLMS algorithm for an IIR system with $M = 3$ and $N = 2$. The system parameters were given by $a_1 = -0.3, a_2 = -0.54, b_0 = 1.0,$ and $b_1 = -2.4$, whereas the settings of the proposed method were identical to the previous experiment. Table III tabulates the means and variances of the BFLMS estimates for four different

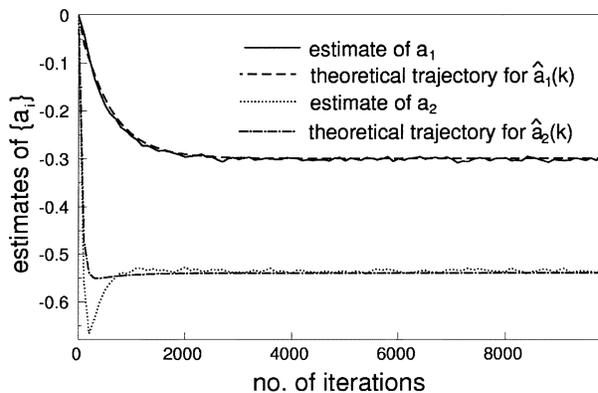


Fig. 7. Estimates of $\{a_i\}$ for BFLMS algorithm in IIR filtering.

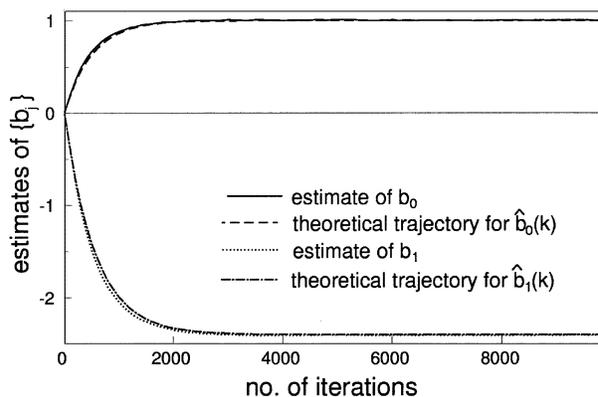


Fig. 8. Estimates of $\{b_j\}$ for BFLMS algorithm in IIR filtering.

noise conditions, and Figs. 7 and 8 show the learning trajectories of $\{\hat{a}_i(k)\}$ and $\{\hat{b}_j(k)\}$ at $\sigma_{n_i}^2 = 0.1$ and $\sigma_{n_o}^2 = 1.0$. From Table III, we see that the proposed approach attained unbiased impulse response estimation and the measured MSEs of the parameter estimates agreed with the predicted values of (15) and (16) for all noise conditions. In addition, the variances of $\{\hat{a}_i(k)\}$ and $\{\hat{b}_j(k)\}$ were smallest for $\sigma_{n_i}^2 = \sigma_{n_o}^2 = 0.1$ and were largest when $\sigma_{n_i}^2 = \sigma_{n_o}^2 = 1.0$. From Figs. 7 and 8, it is observed that $\hat{a}_1(k) \rightarrow -0.3$ and $\hat{b}_0(k) \rightarrow 1.0$ at approximately the 2000th iteration, whereas $\hat{a}_2(k)$ and $\hat{b}_1(k)$ converged to their desired values at the 1000th and the 3000th iteration, respectively. The learning characteristics of the parameter estimates also conformed to their predicted trajectories, except for the transient behavior of $\hat{a}_2(k)$. This discrepancy was due to the assumption of $e^2(k)/(1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k)) \approx \sigma_{n_o}^2$ in deriving (13),

$$\begin{aligned} \text{var}(\hat{a}_i) &\triangleq \lim_{k \rightarrow \infty} E\{(\hat{a}_i(k) - a_i)^2\} \\ &\approx \mu_a \frac{\sigma_{n_o}^2 (r_{dd}(0) + \sigma_{n_o}^2) \left(1 + \sum_{l=1}^{M-1} a_l^2 + \gamma \sum_{l=0}^{N-1} b_l^2\right) - \sigma_{n_o}^4 a_i^2}{2r_{dd}(0)}, \quad i = 1, 2, \dots, M-1 \end{aligned} \quad (15)$$

and

$$\begin{aligned} \text{var}(\hat{b}_j) &\triangleq \lim_{k \rightarrow \infty} E\{(\hat{b}_j(k) - b_j)^2\} \\ &\approx \mu_b \frac{\sigma_{n_o}^2 (r_{ss}(0) + \sigma_{n_i}^2) \left(1 + \sum_{l=1}^{M-1} a_l^2 + \gamma \sum_{l=0}^{N-1} b_l^2\right) - \sigma_{n_i}^4 b_j^2}{2r_{ss}(0)}, \quad j = 0, 1, \dots, N-1 \end{aligned} \quad (16)$$

TABLE III
MEANS AND VARIANCES OF THE PARAMETER ESTIMATES FOR BFLMS ALGORITHM IN IIR FILTERING

Filter Coefficient		a_1	a_2	b_0	b_1
Actual Value		-0.3	-0.54	1.0	-2.4
$\sigma_{n_i}^2 = 0.1$ $\sigma_{n_o}^2 = 1.0$	Mean	-3.021×10^{-1}	-5.364×10^{-1}	1.003×10^0	-2.408×10^0
	Measured MSE	1.669×10^{-3}	1.787×10^{-3}	2.197×10^{-3}	2.032×10^{-3}
	Predicted MSE	2.278×10^{-3}	2.256×10^{-3}	2.253×10^{-3}	2.206×10^{-3}
$\sigma_{n_i}^2 = 0.1$ $\sigma_{n_o}^2 = 0.1$	Mean	-3.008×10^{-1}	-5.399×10^{-1}	1.002×10^0	-2.406×10^0
	Measured MSE	7.915×10^{-4}	7.958×10^{-4}	8.560×10^{-4}	7.790×10^{-4}
	Predicted MSE	8.232×10^{-4}	8.200×10^{-4}	8.856×10^{-4}	8.380×10^{-4}
$\sigma_{n_i}^2 = 1.0$ $\sigma_{n_o}^2 = 1.0$	Mean	-3.090×10^{-1}	-5.379×10^{-1}	1.025×10^0	-2.472×10^0
	Measured MSE	9.269×10^{-3}	9.314×10^{-3}	1.406×10^{-2}	1.467×10^{-2}
	Predicted MSE	9.045×10^{-3}	9.022×10^{-3}	1.528×10^{-2}	1.052×10^{-2}
$\sigma_{n_i}^2 = 1.0$ $\sigma_{n_o}^2 = 0.1$	Mean	-3.070×10^{-1}	-5.415×10^{-1}	1.022×10^0	-2.464×10^0
	Measured MSE	7.579×10^{-3}	7.499×10^{-3}	1.113×10^{-2}	1.138×10^{-2}
	Predicted MSE	6.975×10^{-3}	6.975×10^{-3}	1.379×10^{-2}	8.036×10^{-3}

which possibly could not be applicable to all parameter estimates at the beginning of the adaptation.

V. CONCLUSIONS

An LMS-based algorithm, called the BFLMS method, has been developed for accurate system identification in the presence of input and output interference, given the noise power ratio. The idea is to minimize the mean square value of the equation-error function subject to a constant-norm constraint, and it is proved that the constrained optimization can be converted into an unconstrained optimization of minimizing a scaled mean square error function. Learning behavior and mean square errors of the estimated parameters are derived and verified by computer simulations. The unbiasedness of the proposed approach is demonstrated in both FIR and IIR filtering under different conditions, and it is shown that the BFLMS algorithm outperforms the CAH and TLMS methods in FIR system identification, particularly when the input and output noise powers are not equal.

REFERENCES

- [1] T. Soderstrom and P. Stoica, *System Identification*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [2] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis, Forecasting and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
- [3] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ: Prentice-Hall, 2000.
- [4] S. Sagara and K. Wada, "On-line modified least-squares parameter estimation of linear discrete dynamic systems," *Int. J. Contr.*, vol. 25, no. 3, pp. 329–343, 1977.
- [5] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
- [6] P. A. Regalia, "An unbiased equation error identifier and reduced-order approximation," *IEEE Trans. Signal Processing*, vol. 42, pp. 1397–1412, June 1994.
- [7] K. C. Ho and Y. T. Chan, "Bias removal in equation-error adaptive IIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 51–62, Jan. 1995.
- [8] S. C. Douglas and M. Rupp, "On bias removal and unit norm constraints in equation error adaptive IIR filters," in *Proc. 30th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 1996, pp. 1093–1097.
- [9] H. N. Kim and W. J. Song, "Unbiased equation-error adaptive IIR filtering based on monic normalization," *IEEE Signal Processing Lett.*, vol. 6, pp. 35–37, Feb. 1999.
- [10] C. E. Davila, "An efficient recursive total least squares algorithm for FIR adaptive filtering," *IEEE Trans. Signal Processing*, vol. 42, pp. 268–280, Feb. 1994.
- [11] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM J. Numer. Anal.*, vol. 17, no. 6, pp. 883–893, 1980.
- [12] K. Gao, M. O. Ahmad, and M. N. S. Swamy, "A constrained anti-Herbian learning algorithm for total least-squares estimation with applications to adaptive FIR and IIR filtering," *IEEE Trans. Circuits Syst. II*, vol. 41, pp. 718–729, Nov. 1994.
- [13] D. Z. Feng, Z. Bao, and L. C. Jiao, "Total least mean squares algorithm," *IEEE Trans. Signal Processing*, vol. 46, pp. 2122–2130, Aug. 1998.
- [14] K. V. Frenando and H. Nicholson, "Identification of linear systems with input and output noise: The Koopmans–Levin method," *Proc. Inst. Elect. Eng. D*, vol. 132, no. 1, pp. 30–36, 1985.
- [15] W. X. Zheng, "Adaptive IIR filtering for noisy input-output systems," in *Proc. Int. Symp. Signal Process. Applicat.*, Aug. 1999, pp. 375–378.