Statistical Properties of the LMS Fourier Analyzer in the Presence of Frequency Mismatch

Yegui Xiao, Member, IEEE, Akira Ikuta, Liying Ma, Li Xu, Member, IEEE, and Rabab Kreidieh Ward, Fellow, IEEE

Abstract—The statistical performances of the conventional adaptive Fourier analyzers, such as the least mean square (LMS), the recursive least square (RLS) algorithms, and so forth, may degenerate significantly, if the signal frequencies given to the analyzers are different from the true signal frequencies. This difference is referred to as frequency mismatch (FM). In this paper, we analyze extensively the performance of the conventional LMS Fourier analyzer in the presence of FM. Difference equations governing the dynamics and closed-form steady-state expression for the estimation mean square error (MSE) of the algorithm are derived in detail. It is revealed that the discrete Fourier coefficient (DFC) estimation problem in the LMS eventually reduces to a DFC tracking one due to the FM, and an additional term derived from DFC tracking appears in the closed-form MSE expression, which essentially deteriorates the performance of the algorithm. How to derive the optimum step size parameters that minimize or mitigate the influence of the FM is also presented, which can be used to perform robust design of step size parameters for the LMS algorithm in the presence of FM. Extensive simulations are conducted to reveal the validity of the analytical results.

Index Terms—Adaptive Fourier analyzer, convergence, frequency mismatch (FM), mean square error (MSE), optimum step size parameters, performance analysis, sinusoidal signal.

I. INTRODUCTION

THERE are many applications in digital communications, power systems, control engineering, active noise and/or vibration control, biomedical engineering, and pitch detection in automated transcription, just to name a few, where we are concerned with the analysis of a sinusoidal signal in additive noise [1]–[9], [12], [14]. The frequencies of the sinusoidal signal are arbitrary, and are assumed to be known or estimated in advance. Furthermore, the signal is nonstationary for most of the time. The discrete Fourier transform (DFT) and its variants may be used to perform robust design of step size parameters for the LMS algorithm in the presence of FM. Extensive simulations are conducted to reveal the validity of the analytical results.

Manuscript received June 13, 2003; revised March 18, 2004. This paper was presented in part at the IEEE International Symposium on Circuits and Systems (ISCAS2003), Bangkok, Thailand, 25–28 May, 2003. This work was supported in part by The Research Foundation for Electrotechnology of Chubu (REFEC), Nagoya, Japan under Grant R-12215. This paper was recommended by Associate Editor W. X. Zheng.

Y. Xiao and A. Ikuta are with the Faculty of Human Life and Environmental Science, Hiroshima Prefectural Women’s University (HPWU), Hiroshima 734-8558, Japan (e-mail: xiao@hirojo-u.ac.jp).

L. Ma is with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC H3G 1M8, Canada.

L. Xu is with the Department of Electronic Information Systems, Akita Prefectural University, Akita 015-0055, Japan.

R. Kreidieh Ward is with the Institute for Computing, Information and Cognitive Systems, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.

Digital Object Identifier 10.1109/TCSI.2004.838315
the statistical frequency estimation methods are usually biased when applied to a real-life data set of limited length. The behavior of an FIR-type LMS active noise canceling system with sinusoidal reference input in the presence of frequency differences (or FM) was discussed qualitatively in [19], but no quantitative analysis was provided. An approximated MSE expression for the LMS Fourier analyzer (linear combiner) in the presence of FM was given in [13] based on the analysis in the frequency domain, but detailed statistical analysis was not performed. The backgrounds regarding the existence of the FM in real-life applications have not been provided in both [19] and [13].

In this paper, first, we show, by simulations, the performance degradation of the LMS algorithm due to the FM. Next, we analyze the influence of the FM on the LMS algorithm when it is used to analyze a multifrequency sinusoidal signal in additive noise. This analysis makes it easier to understand the essence of the problem. As a result, the dynamics of the LMS with FM are derived, and its steady-state statistical properties are also evaluated. As a by-product, optimum step size parameters are also derived which minimize the total MSE of the algorithm.

The rest of the paper is organized as follows. Section II introduces the conventional LMS algorithm, and uncovers its problem with the FM by showing a typical simulation result. The behaviors of the LMS algorithm in the presence of FM are analyzed in detail in Section III. In Section IV simulation results are provided to support the analytical findings. Conclusions are given in Section V.

II. CONVENTIONAL LMS FOURIER ANALYZER

Here, assume a multifrequency sinusoidal signal in additive noise

\[ d(n) = s(n) + v(n) \]

\[ = \sum_{i=1}^{q} \{ a_i \cos(\omega_{0i} n) + b_i \sin(\omega_{0i} n) \} + v(n) \]

where \( n \) is the discrete time instant, \( q \) is the known number of frequency components of the sinusoidal signal \( s(n) \), \( \omega_{0i} \) is the frequency of the \( i \)th component of the signal to be analyzed, and \( v(n) \) is a zero-mean additive white noise with variance \( \sigma_v^2 \). The frequencies \( \{ \omega_{0i} \}_{i=1}^{q} \) are assumed to be given or known in advance. Fig. 1 shows the conventional real-valued LMS algorithm [13], [14] which is derived from the complex-valued LMS algorithm [6], [7]. The purpose of the LMS Fourier analyzer is to estimate, in real time, the discrete Fourier coefficients (DFCs) \( \{ a_i, b_i \}_{i=1}^{q} \) by minimizing the mean squared error \( E[e^2(n)] \) with respect to \( \{ \hat{a}_i(n), \hat{b}_i(n) \}_{i=1}^{q} \). The recursion of the (real-valued) LMS algorithm is given by

\[ \hat{D}_i(n+1) = \hat{D}_i(n) + \mu_i e(n)X_i(n), \quad i = 1, 2, \cdots, q \]

where

\[ \hat{D}_i(n) = [\hat{a}_i(n), \hat{b}_i(n)]^T \]

\[ X_i(n) = [\cos(\omega_{in}), \sin(\omega_{in})]^T \]

\[ e(n) = d(n) - y(n) \]

\[ y(n) = \sum_{i=1}^{q} \hat{D}_i(n)^T X_i(n). \]

Here, \( \mu_i \) is the step size parameter for the DFCs of the \( i \)th component, which controls the magnitude of the recursion, and \( \omega_{i} \) is a user-specified frequency for the \( i \)th component. \( \omega_{0i} \) is supposed to be the same as the real frequency \( \omega_{0i} \). However, as mentioned in Section I, in real applications, there may be a mismatch between \( \omega_{0i} \) and \( \omega_{i} \). That is, the FM

\[ \Delta \omega_i = \omega_{0i} - \omega_{i}, \quad i = 1, 2, \cdots, q \]

may be not zero. Extensive simulations have revealed that when the FM vector \( \Delta \Omega = [\Delta \omega_1, \Delta \omega_2, \cdots, \Delta \omega_q]^T \) corresponding to the signal frequency vector \( \Omega = [\omega_1, \omega_2, \cdots, \omega_q]^T \) is not a very small vector, the performance of the LMS may degrade significantly. To show the influence of the FM, an example is provided in Fig. 2. It is obviously seen that 1% of FM gives rise to a significant performance degradation. This implies that the conventional LMS is not very robust to the FM. We have also found that other algorithms, such as the RLS, the SRLS, the Kalman filtering algorithms and so on, also present similar performance degradation in the presence of FM, even though we know that in several applications, the LMS-type algorithms have been found effective and efficient [7], [9], [12], [14]. Addressing first how the LMS algorithm behaves in the presence of FM will definitely enrich our understanding of its properties.
and stimulate further research for better and more robust adaptive algorithms. This work is devoted to showing quantitatively how vulnerable the LMS algorithm is when confronted with the existence of FM, and searching for possible ways to reduce the performance degeneration.

III. ANALYSIS OF LMS ALGORITHM IN THE PRESENCE OF FM

In this section, we provide a detailed performance analysis of the LMS algorithm in the presence of FM.

A. Error Signal

Substituting (2) and (7) into (6), we have, after some manipulations

\[
e(n) = \sum_{i=1}^{q} \left( \delta_{a_i}(n) \cos(\omega_i n) + \delta_{b_i}(n) \sin(\omega_i n) \right) + v(n)
\]

(9)

where the estimation errors for the DFCs are defined by

\[
\delta_{a_i}(n) = a_i \cos(\Delta \omega_i n) + b_i \sin(\Delta \omega_i n) - \hat{a}_i(n)
\]

\[
\delta_{b_i}(n) = b_i \cos(\Delta \omega_i n) - a_i \sin(\Delta \omega_i n) - \hat{b}_i(n).
\]

(10, 11)

Clearly, due to the existence of the FM, the DFC estimates \( \hat{a}_i(n) \) and \( \hat{b}_i(n) \) will try to track the time-varying signals \( a_i \cos(\Delta \omega_i n) + b_i \sin(\Delta \omega_i n) \) and \( b_i \cos(\Delta \omega_i n) - a_i \sin(\Delta \omega_i n) \), respectively, rather than the constants \( a_i \) and \( b_i \), in order to achieve minimum MSE \( E[\varepsilon^2(n)] \). This implies that for the LMS algorithm, the estimation problem in the presence of FM eventually becomes a tracking problem which is generally more difficult.

B. Convergence and Steady-State Analysis of the Tracking Errors

Substituting (9)–(11) into (3) yields

\[
\delta_{a_i}(n + 1) = \delta_{a_i}(n) - \mu_i \left\{ \sum_{j=1}^{q} \left( \delta_{a_j}(n) \cos(\omega_j n) + \delta_{b_j}(n) \sin(\omega_j n) \right) \right\} \cos(\omega_i n) + f_{a_i}(n)
\]

(12)

where

\[
f_{a_i}(n) = - \xi_i \sin(\Delta \omega_i n) + \eta_i \cos(\Delta \omega_i n)
\]

(13)

\[
\xi_i = a_i \sin(\Delta \omega_i) + b_i (1 - \cos(\Delta \omega_i))
\]

(14)

\[
\eta_i = b_i \sin(\Delta \omega_i) - a_i (1 - \cos(\Delta \omega_i)).
\]

(15)

The ensemble-averaging equation (12) leads to

\[
E[\delta_{a_i}(n + 1)] = (1 - \mu_i E[\cos^2(\omega_i n)]) E[\delta_{a_i}(n)]
\]

\[\quad - \mu_i \sum_{j=1,j\neq i}^{q} E[\delta_{a_j}(n)] E[\cos(\omega_j n) \cos(\omega_j n)]
\]

\[\quad - \mu_i \sum_{j=1}^{q} E[\delta_{b_j}(n)] E[\cos(\omega_j n) \sin(\omega_j n)] + f_{a_i}(n)
\]

\[\quad = \left( 1 - \frac{1}{2} \mu_i \right) E[\delta_{a_i}(n)] + f_{a_i}(n).
\]

(16)

In the above derivations, we treated \( f_{a_i}(n) \) as a deterministic signal, because it has a period which is much longer than that of the sinusoidal signal \( s(n) \), i.e., \( \Delta \omega_i \ll \omega_i \) for all \( i \). Moreover, \( \sin(\omega_i n) \) and \( \cos(\omega_j n) \) and their functions, such as \( \cos^2(\omega_i n) \), \( \cos(\omega_i n) \sin(\omega_j n) \) and so on, are treated as pseudorandom signals whose ensemble averaging is replaced by their time averaging [22, 23].

Similarly, for the estimation error of DFC of the sine component, one can show that

\[
E[\delta_{b_i}(n + 1)] = \left( 1 - \frac{1}{2} \mu_i \right) E[\delta_{b_i}(n)] + f_{b_i}(n)
\]

(17)

where

\[
f_{b_i}(n) = - \eta_i \sin(\Delta \omega_i n) - \xi_i \cos(\Delta \omega_i n).
\]

(18)

Here, the following comments are in order.

CB1 It can be seen from (16) and (17) that the conventional LMS algorithm is no longer unbiased due to the existence of the FM, and the estimation of DFCs essentially becomes a problem of tracking two sinusoidal components. That is, the estimation errors are actually equal to the tracking errors which fluctuate with a frequency equal to the FM at the steady state of the algorithm.

CB2 In spite of the existence of the FM, the convergence of the LMS algorithm in the mean is still uniform. Furthermore, the larger the step size value (the smaller the convergence time constant) within the stability bound of the algorithm, the faster the algorithm will reach its steady state in the sense of convergence of the mean.

CB3 For an extremely small FM, one gets \( \xi_i \approx a_i \Delta \omega_i \) and \( \eta_i \approx b_i \Delta \omega_i \), which implies that the amplitudes of \( f_{a_i}(n) \) and \( f_{b_i}(n) \) are also very small. Therefore the fluctuations of the tracking errors are approximately proportional to the FM, and are very small as well. This implies that the conventional LMS is able to work sufficiently well if the FM is tiny. See [7], [9], [12], [14] for the past successful applications.

CB4 If the FM is zero, the above difference equations are identical to those obtained in [13], [14].

To study the steady-state behaviors of the tracking errors \( E[\delta_{a_i}(n)] \) and \( E[\delta_{b_i}(n)] \), we take the z-transforms of (16) and (13). This eventually leads to

\[
Z \{ E[\delta_{a_i}(n)] \} = \frac{z}{z - \alpha_i} \frac{\eta_i z + \gamma_i}{z^2 - 2z \cos(\Delta \omega_i) + 1}
\]

(19)

where

\[
\alpha_i = 1 - \frac{1}{2} \mu_i
\]

(20)

\[
\gamma_i = - \xi_i \sin(\Delta \omega_i) - \eta_i \cos(\Delta \omega_i).
\]

(21)

Taking the inverse z-transform of (19) gives

\[
E[\delta_{a_i}(n)] = c_{a_i,1} \alpha_i^n + c_{a_i,2} \cos(\Delta \omega_i n)
\]

\[\quad + c_{a_i,3} \frac{\cos(\Delta \omega_i n) + \xi_i}{\sin(\Delta \omega_i)} \sin(\Delta \omega_i n)
\]

(22)
where

\[ c_{i,1} = \frac{\alpha \eta_i + \gamma_i}{\alpha_i^2 - 2 \alpha_i \cos \Delta \omega_i + 1}, \quad (23) \]

\[ c_{i,2} = c_{i,3}, \quad (24) \]

\[ c_{i,3} = \frac{c_{i,1} - \gamma_i}{\alpha_i}. \quad (25) \]

Similarly, one has, from (17) and (18)

\[ E[\delta_i(n)] = -d_{i,1} \alpha_i^2 - d_{i,2} \cos(\Delta \omega_i n) - \frac{d_{i,3} \cos \Delta \omega_i + d_{i,3}}{\sin \Delta \omega_i} \sin(\Delta \omega_i n) \quad (26) \]

where \( d_{i,1}, d_{i,2}, \) and \( d_{i,3} \) satisfy the same (23)–(25) for \( c_{i,1}, c_{i,2}, \) and \( c_{i,3} \), if one replaces \( \eta_i \) and \( \gamma_i \) in (23), (25) by \( \xi_i \) and \( \eta_i \sin \Delta \omega_i - \xi_i \cos \Delta \omega_i \), respectively.

At steady state, it can be noticed that the tracking errors \( E[\delta_i(n)] \) and \( E[\delta_i(n)] \), fluctuate sinusoidally with a frequency that is exactly the same as the FM. This indicates that the tracking errors are very slowly time-varying compared with the sinusoid being analyzed. Furthermore, if \( \alpha_i < 1 \) or \( 0 < \mu_i < 2 \), the LMS algorithm is guaranteed to converge in the mean. Note that this also can be deduced by examining the difference equations (16) and (17). It is worthwhile to note that the z-transform based technique used to analyze the above difference equations (16) and (17) has been utilized to analyze the properties of an FIR-type adaptive noise canceling system and the conventional LMS Fourier analyzer in the frequency domain in [19] and [13], respectively. The subject of this subsection is discussed in [24] in similar manner in the context of adaptive suppression of tremor for improved human-machine control.

At steady state, the first terms in the RHS of (22) and (26) will vanish and the tracking errors, \( E[\delta_i(n)] \) and \( E[\delta_i(n)] \), reduce to sinusoidal waves with a frequency that is exactly the FM. The amplitudes of these sine waves indicate the magnitudes of the tracking errors. At a glance, they can be easily calculated from (22) and (26), but it is not difficult to find out, by a close inspection, that they are very complicated functions of the FM, the DFCs of the sinusoids and the step size parameters. Numerical calculations are needed in order to explain the relationships among the tracking errors, the FM and the step size parameters. Extensive numerical calculations show that: 1) for the same FM (or the same time constant of FM), a larger step size value (or a smaller convergence time constant) results in smaller tracking errors; 2) for the same convergence time constant or step size value, a larger FM or a smaller time constant of the FM leads to larger tracking errors. This suggests that one needs to use a larger step size value in order to achieve smaller tracking errors and MSEs due to a larger FM. However, the larger step size values will raise the tracking MSEs due to the additive noise. Therefore, a wise tradeoff is inevitable in selecting the step size values in order to achieve the desired performance. See Section III-D for details.

C. Convergence and Steady-State Analysis of the Tracking MSEs

Squaring (12) and ensemble-averaging the resultant MSE equation, after complicated calculations, one yields

\[
E[\delta_i^2(n+1)] = \left( 1 - \mu_i + \frac{3}{8} \mu_i^2 \right) E[\delta_i^2(n)] + \frac{1}{8} \mu_i^2 E[\delta_i^2(n)] + \frac{1}{4} \mu_i^2 \sum_{j=1}^{q} \left( E[\delta_{ij}^2(n)] + E[\delta_{ij}^2(n)] \right)
\]

\[
+ \frac{1}{8} \mu_i^2 \sum_{j=1}^{q} \sum_{l=1, l \neq j}^{q} E[\delta_{ij}^2(n)] E[\delta_{ij}^2(n)] \times \{ \delta(\omega_{ij} + \omega_{jl} - 2 \omega_i) + \delta(\omega_{ij} - \omega_{jl} - 2 \omega_i) \}
\]

\[
- \frac{1}{8} \mu_i^2 \sum_{j=1}^{q} \sum_{l=1, l \neq j}^{q} E[\delta_{ij}^2(n)] E[\delta_{ij}^2(n)] \times \{ \delta(\omega_{ij} + \omega_{jl} - 2 \omega_i) - \delta(\omega_{ij} - \omega_{jl} - 2 \omega_i) \}
\]

\[
+ (2 - \mu_i)f_i(n)E[\delta_{i}^2(n)] + f_i^2(n) + \frac{1}{2} \mu_i^2 \sigma_v^2 \quad (27)
\]

where \( \delta(\cdot) \) is a Dirac delta function. In the same way, one reaches

\[
E[\delta_i^2(n+1)] = \frac{1}{8} \mu_i^2 E[\delta_i^2(n)] + \left( 1 - \mu_i + \frac{3}{8} \mu_i^2 \right) E[\delta_i^2(n)]
\]

\[
+ \frac{1}{4} \mu_i^2 \sum_{j=1}^{q} \left( E[\delta_{ij}^2(n)] + E[\delta_{ij}^2(n)] \right)
\]

\[
- \frac{1}{8} \mu_i^2 \sum_{j=1}^{q} \sum_{l=1, l \neq j}^{q} E[\delta_{ij}^2(n)] E[\delta_{ij}^2(n)] \times \{ \delta(\omega_{ij} + \omega_{jl} - 2 \omega_i) + \delta(\omega_{ij} - \omega_{jl} - 2 \omega_i) \}
\]

\[
+ \frac{1}{8} \mu_i^2 \sum_{j=1}^{q} \sum_{l=1, l \neq j}^{q} E[\delta_{ij}^2(n)] E[\delta_{ij}^2(n)] \times \{ \delta(\omega_{ij} + \omega_{jl} - 2 \omega_i) - \delta(\omega_{ij} - \omega_{jl} - 2 \omega_i) \}
\]

\[
+ (2 - \mu_i)f_i(n)E[\delta_i^2(n)] + f_i^2(n) + \frac{1}{2} \mu_i^2 \sigma_v^2 \quad (28)
\]

From the above two difference equations, we see that

CC1) Due to the FM, the difference equations (27), (28) are correlated with the tracking errors \( \delta_i(n) \) and \( \delta_i(n) \). Therefore, to numerically study the dynamics of the LMS, difference equations for the convergences in the mean (16), (17) and mean square (27), (28) have to be solved simultaneously.

CC2) The steady-state MSEs, \( E[\delta_i^2(n)] \) and \( E[\delta_i^2(n)] \), indicate fluctuations with a frequency that is just twice of the FM. This is because i) the fourth and fifth terms in the right hand sides (RHS’s) of (27) and (28) are usually very small, when compared with other terms, and therefore (27) and (28) are almost linear with respect to the estimation MSEs \( E[\delta_i^2(n)] \) and \( E[\delta_i^2(n)] \); ii) the sixth and seventh terms in the RHS of (27) and (28) present a
sinusoidal signal whose frequency is just twice of the FM.

CC3) Basically, from the four difference equations (16), (17) and (27), (28) for the convergences in the mean and in the mean square, stability bounds for the step size parameters $\mu_i (i = 1, 2, \ldots, q)$ may be obtained. It seems that analytical derivation of these bounds is complicated and difficult, but numerical solutions to them can be easily obtained based on a grid search (see also Section IV).

CC4) If the FM vanishes and one ignores the fourth and fifth terms in the RHS of (27) and (28), these difference equations reduce to those obtained in [13], [14]. Note that these two terms are neglected in [13], [14]. It can be noted from (27), (28) that the tracking MSEs fluctuate even at their steady states due to the existence of the FM, just like the tracking errors do. Here, we further introduce time averaging, indicated by $E_T[\cdot]$, to the ensemble-averaged tracking MSEs to simplify the analysis. Before we move to the derivation of the time-averaged steady-state MSEs, we make some useful definitions. From (9), one easily gets

$$E_T [\epsilon^2(n)] = \frac{1}{2} \sum_{i=1}^{q} \{ E_T [\delta^2_{a_i}(n)] + E_T [\delta^2_{b_i}(n)] \} + \sigma_v^2.$$  

(29)

Now define

$$J_i(n) = E_T [\delta^2_{a_i}(n)] + E_T [\delta^2_{b_i}(n)].$$

(30)

$$J(n) = \sum_{i=1}^{q} J_i(n).$$

(31)

After imposing time averaging, one may assume

$$E_T [\delta^2_{a_i}(n+1)] \big|_{n=\infty} = E_T [\delta^2_{a_i}(n)] \big|_{n=\infty} = E_T [\delta^2_{a_i}(\infty)],$$

$$E_T [\delta^2_{b_i}(n+1)] \big|_{n=\infty} = E_T [\delta^2_{b_i}(n)] \big|_{n=\infty} = E_T [\delta^2_{b_i}(\infty)].$$

Taking the time averaging of (27) and (28) at the steady state and using the above relations, one yields

$$E_T [\delta^2_{a_i}(\infty)] \approx \left(1 - \mu_i + \frac{3}{8} \mu_i^2 \right) E_T [\delta^2_{a_i}(\infty)] + \frac{1}{4} \mu_i^2 \sum_{j=1, j \neq i}^{q} \left( E_T [\delta^2_{a_j}(\infty)] + E_T [\delta^2_{b_j}(\infty)] \right)$$

$$+ \frac{1}{2} \mu_i^2 \left( F_i(\mu_i, \Delta \omega_i) + (1 - \cos \Delta \omega_i) A_i^2 \right) + \sigma_v^2.$$  

(32)

$$E_T [\delta^2_{b_i}(\infty)] \approx \frac{1}{8} \mu_i^2 E_T [\delta^2_{a_i}(\infty)] + \left(1 - \mu_i + \frac{3}{8} \mu_i^2 \right) E_T [\delta^2_{b_i}(\infty)]$$

$$+ \frac{1}{4} \mu_i^2 \sum_{j=1, j \neq i}^{q} \left( E_T [\delta^2_{a_j}(\infty)] + E_T [\delta^2_{b_j}(\infty)] \right)$$

$$+ \frac{1}{2} \mu_i^2 \left( F_i(\mu_i, \Delta \omega_i) + (1 - \cos \Delta \omega_i) A_i^2 \right) + \sigma_v^2.$$  

(33)

where

$$F_i(\mu_i, \Delta \omega_i) = F_i(\mu_i, \Delta \omega_i)$$

$$= \frac{1}{\beta_i} \left( \alpha_i - \cos \Delta \omega_i \right) \left( 1 - \cos \Delta \omega_i \right) A_i^2$$

(34)

$$\beta_i = \alpha_i^2 - 2 \alpha_i \cos \Delta \omega_i + 1.$$  

(35)

$$A_i^2 = \alpha_i^2 + \beta_i.$$  

(36)

By subtracting (33) from (32), one easily reaches

$$\left( \mu_i - \frac{1}{4} \mu_i^2 \right) \left( E_T [\delta^2_{a_i}(\infty)] - E_T [\delta^2_{b_i}(\infty)] \right) = 0$$

which implies

$$E_T [\delta^2_{a_i}(\infty)] = E_T [\delta^2_{b_i}(\infty)] = \frac{1}{2} J_i(n) \big|_{n=\infty} = \frac{1}{2} J_i(\infty).$$

(37)

Adding (32) and (33) readily leads to

$$\left( \mu_i - \frac{1}{2} \mu_i^2 \right) J_i(\infty) = \frac{1}{2} \mu_i^2 \sum_{j=1, j \neq i}^{q} J_j(\infty) = 2 F_i(\mu_i, \Delta \omega_i).$$

(38)

One can obtain from (38)

$$J_i(\infty) = \frac{1}{2} \mu_i J_i(\infty) + \frac{2}{\mu_i} F_i(\mu_i, \Delta \omega_i)$$

$$= \frac{\sum_{j=1}^{q} \frac{1}{\mu_j} F_j(\mu_j, \Delta \omega_j)}{2 - \sum_{j=1}^{q} \frac{1}{\mu_j}} + \frac{2}{\mu_i} F_i(\mu_i, \Delta \omega_i).$$

(39)

As a result, the time-averaged MSE turns out to be

$$E_T [\epsilon^2(n)] \big|_{n=\infty} = \frac{1}{2} J(n) + \sigma_v^2 \bigg|_{n=\infty} = \frac{1}{2} \mu_i J(n) \big|_{n=\infty} + \sigma_v^2.$$  

(40)

$$= \frac{1}{2} \mu_i J(n) \big|_{n=\infty} + \sigma_v^2.$$  

(42)

D. Minimization of the Steady-State Tracking MSEs

Here, for a relatively small FM, we first introduce the following approximations that are required to make the minimization analytically simple. They are based on the assumptions of relatively slow adaptation (small step size parameters) and small
FM, and also on the use of Taylor expansion of $\cos \Delta \omega_j$ in the vicinity of zero up to the second-order term.

$$\cos \Delta \omega_j \approx 1 - \frac{1}{2} \Delta \omega_j^2$$  \hspace{1cm} (43)

$$\alpha_j - \cos \Delta \omega_j \approx 1 - \frac{\mu_j}{2} - \left(1 - \frac{1}{2} \Delta \omega_j^2\right)$$

$$= \frac{1}{2} \left(\Delta \omega_j^2 - \mu_j\right)$$  \hspace{1cm} (44)

$$\beta_j \approx \left(1 - \frac{\mu_j}{2}\right)^2 - 2 \left(1 - \frac{\mu_j}{2}\right) \left(1 - \frac{1}{2} \Delta \omega_j^2\right) + 1$$

$$\approx \frac{1}{4} \mu_j^2 + \Delta \omega_j^2.$$  \hspace{1cm} (45)

Using the above approximations in $F_{a_j}(\mu_j, \Delta \omega_j)$ readily yields

$$F_{a_j}(\mu_j, \Delta \omega_j) \approx - \frac{\Delta \omega_j^2 A_j^2}{\mu_j^2 + 4 \Delta \omega_j^2} \left(\Delta \omega_j^2 - \mu_j\right).$$  \hspace{1cm} (46)

Then, one gets

$$\frac{2}{\mu_j} F_{j}(\mu_j, \Delta \omega_j) = \left\{ \frac{2 - \mu_j}{\mu_j} F_{a_j}(\mu_j, \Delta \omega_j) \right\}$$

$$+ \frac{1}{2} \mu_j \sigma_v^2 + \frac{1}{\mu_j} \left(1 - \cos \Delta \omega_j \right) \sigma_v^2$$

$$\approx 2 \left(\frac{1}{2} \mu_j \sigma_v^2 + \frac{2 \Delta \omega_j^2 A_j^2}{\mu_j^2 + 4 \Delta \omega_j^2}\right).$$  \hspace{1cm} (47)

Now, we have the following two steps to minimize the $E_T[\varepsilon^2(n)]\big|_{n=\infty}$ (42)) with respect to the step size parameters.

Step 1) Minimize $\left(\frac{2}{\mu_j} F_{j}(\mu_j, \Delta \omega_j)\right)$ for $j = 1, 2, \cdots, q$. This requires

$$\frac{\partial}{\partial \mu_j} \left(\frac{2}{\mu_j} F_{j}(\mu_j, \Delta \omega_j)\right) = \sigma_v^2 - \frac{8 \mu_j \Delta \omega_j^2 A_j^2}{\left(\mu_j^2 + 4 \Delta \omega_j^2\right)^2} = 0$$

which approximately produces a cubic equation in the step size parameter $\mu_j$

$$\mu_j^3 + 8 \Delta \omega_j^2 \mu_j - \frac{8 \Delta \omega_j^2 A_j^2}{\sigma_v^2} = 0.$$  \hspace{1cm} (48)

Based on the Cardano’s formula, this cubic equation only has one real solution that is approximated and given by

$$\mu_j^{(1)}_{opt} \approx 2 \sqrt[3]{\frac{\Delta \omega_j^2 A_j^2}{\sigma_v^2}}.$$  \hspace{1cm} (49)

Step 2) Use the steepest-descent algorithm to minimize the time-averaged MSE (42) with respect to all the step size parameters, while taking $\mu_j^{(1)}_{opt}$ ($j = 1, 2, \cdots, q$) as their initial values.

It should be noted that the above results from the MSE minimization can be easily and directly applied to calculate the step size parameters if the LMS algorithm is to be used to detect the signal, for example, in DTMF signaling, where the maximum FM and the signal and noise powers are known or approximated in advance.

### E. Steady-State Estimation Bias of the Squared Amplitude

In the detection of sinusoids with frequencies of interests, the estimated amplitude or squared amplitude of each sinusoid can be used. Therefore, it is useful to investigate how the FM affects the amplitude estimation of the LMS algorithm. We here investigate $b_r^2(n) + b_i^2(n)$, rather than the amplitude $\sqrt{b_r^2(n) + b_i^2(n)}$, because the former is easier to handle.

From (10) and (11), we have

$$\hat{A}_r^2(n) = \hat{a}_r^2(n) + \hat{b}_i^2(n)$$

$$= \hat{a}_r^2 + \hat{b}_i^2 + \hat{a}_r(n) \hat{a}_r(n) + \hat{b}_i(n) \hat{b}_i(n)$$

$$- 2 \hat{a}_r(n) \hat{a}_i(n) g_{a_i}(n) - 2 \hat{b}_i(n) g_{b_i}(n)$$  \hspace{1cm} (50)

where

$$g_{a_i}(n) = a_i \cos(\Delta \omega_i n) + b_i \sin(\Delta \omega_i n)$$  \hspace{1cm} (51)

$$g_{b_i}(n) = b_i \cos(\Delta \omega_i n) - a_i \sin(\Delta \omega_i n).$$  \hspace{1cm} (52)

After time averaging the ensemble-averaged equation (50), we get the steady-state bias for the squared amplitude estimate

$$E_T \left[ \delta_{A_r^2}(n) \right]_{n=\infty}$$

$$= E_T \left[ \delta_{A_r^2}(\infty) \right] = E_T \left[ \hat{A}_r^2(n) \right]_{n=\infty} - A_r^2$$

$$= \left\{ E_T \left[ \delta_{a_r^2}(n) \right] + E_T \left[ \delta_{b_i^2}(n) \right] \right\}$$

$$- 2 E_T \left[ E \left[ b_i(n) g_{a_i}(n) \right] \right]_{n=\infty}$$

$$- 2 E_T \left[ E \left[ b_i(n) g_{b_i}(n) \right] \right]_{n=\infty}.$$  \hspace{1cm} (53)

At steady state, after some complicated calculations, we have

$$E_T \left[ E \left[ \delta_{a_i}(n) g_{a_i}(n) \right] \right]_{n=\infty}$$

$$= E_T \left[ E \left[ b_i(n) g_{b_i}(n) \right] \right]_{n=\infty}$$

$$= \frac{1}{2 \mu_i} (\alpha_i + 1)(1 - \cos \Delta \omega_i) A_r^2.$$  \hspace{1cm} (54)

Then, we have

$$E_T \left[ \delta_{A_r^2}(\infty) \right] = J_i(\infty) - \frac{2(\alpha_i + 1)}{\beta_i} (1 - \cos \Delta \omega_i) A_r^2.$$  \hspace{1cm} (55)

Now, the following comments are in order.

CE1) From (55), we see that the LMS algorithm produces a biased estimate for the squared amplitude even when the FM does not exist and/or the true amplitude is zero. The bias in such cases is given by $J_i(\infty)$. If there is no FM, the squared amplitude threshold for detection must be at least larger than this bias. Therefore, (55) can help the user in selecting the squared amplitude threshold for the detection task.

CE2) The biases in the estimated amplitude and the squared amplitude can be positive or negative. If the FM is large, the magnitudes of these biases may become quite large. Interestingly, when they are positive, they favor the detection of a sinusoid that does exist, and do not favor the detection when it does not exist, as the estimated amplitude is larger than the true one on the average. The opposite happens when they are negative. They do not favor the detection of a sinusoid that does
exist, and favor the detection when it does not exist, because the estimated amplitude tends to be smaller than the true one. Thus, every effort should be made to avoid the latter case. Furthermore, these biases are basically harmful in terms of signal estimation accuracy. Therefore, on the whole, reducing the portion of these biases due to the existence of the FM will be beneficial for both signal detection and estimation.

CE3) From (55), we have found by numerical analysis that, as the FM gets larger in magnitude, the magnitudes of the biases of the estimated amplitude and the squared amplitude also become larger. This will degenerate the
F. Steady-State Phase Estimation Bias

For completeness of analysis, an approximate steady-state expression for the phase estimation bias is also derived. From (10) and (11), we have

$$\frac{\hat{b}_i(n)}{a_i(n)} = \frac{b_i \cos(\Delta \omega_i n) - a_i \sin(\Delta \omega_i n) - \delta_{b_i}(n)}{a_i \cos(\Delta \omega_i n) + b_i \sin(\Delta \omega_i n) - \delta_{a_i}(n)}$$

$$\approx \frac{b_i}{a_i} \frac{(a_i^2 + b_i^2) \sin(\Delta \omega_i n) - b_i \delta_{a_i}(n) + a_i \delta_{b_i}(n)}{a_i^2 \cos(\Delta \omega_i n) + b_i^2 \sin(\Delta \omega_i n)}. \quad (56)$$

Then, the phase estimate can be approximately expressed as follows by using of the first-order Taylor series expansion:

$$\hat{\theta}_i(n) = \arctan \left( \frac{\hat{b}_i(n)}{a_i(n)} \right) \approx \theta_i - \frac{\hat{b}_i(n)}{a_i(n)}$$

$$\approx \theta_i - \frac{b_i}{a_i} \frac{(a_i^2 + b_i^2) \sin(\Delta \omega_i n) - b_i \delta_{a_i}(n) + a_i \delta_{b_i}(n)}{a_i^2 \cos(\Delta \omega_i n) + b_i^2 \sin(\Delta \omega_i n)}. \quad (57)$$

The phase estimation bias is given by

$$E \left[ \hat{\theta}_i(n) \right] - \theta_i = - \frac{a_i \sin(\Delta \omega_i n)}{a_i \cos(\Delta \omega_i n) + b_i \sin(\Delta \omega_i n)}$$

$$\times \left( -b_i E \left[ \delta_{a_i}(n) \right] + a_i E \left[ \delta_{b_i}(n) \right] \right). \quad (58)$$

It can be clearly seen that (i) the bias is time-varying even when the algorithm reaches its steady state, and (ii) no bias exists if the FM is zero. If one first uses (22) and (26) into (58) and then takes the time averaging of the resulting equation, the steady-state phase bias turns out to be

$$E_T \left[ \hat{\theta}_i(n) \right] \big|_{n \to \infty} - \theta_i$$

$$= \left\{ 1 - \frac{(b_i c_{i,2} + a_i d_{i,2}) \cos \Delta \omega_i + b_i c_{i,3} + a_i d_{i,3}}{(a_i^2 + b_i^2) \sin \Delta \omega_i} \right\} I_1(i)$$

$$+ \frac{b_i c_{i,2} + a_i d_{i,2}}{a_i^2 + b_i^2} I_2(i), \quad (59)$$

where

$$I_1(i) = \frac{1}{2 \pi} \int_{-\pi}^{\pi} \frac{a_i \sin x}{a_i \cos x + b_i \sin x} dx$$

$$I_2(i) = \frac{1}{2 \pi} \int_{-\pi}^{\pi} \frac{a_i \cos x}{a_i \cos x + b_i \sin x} dx. \quad (60)$$

Fig. 3. (Continued) (c) Third frequency component.
IV. SIMULATION RESULTS

The representative simulation results shown in this Section demonstrate the validity of the analytical results we have obtained so far.

A. Dynamics

First, the theoretical difference equations for the convergences in the mean and mean square senses are compared with their simulated points in Fig. 3, where excellent agreements
between the difference equations and the simulated dynamics can be clearly observed. In this simulation all the frequency components have equal power. The theoretical curves are obtained by solving (16), (17) and (27), (28) as simultaneous difference equations. A simulation for unequal-power case is also performed, where the power of one sinusoid is much stronger than ones of other sinusoids. The comparisons between theory and simulation for the strongest and the weakest sinusoid are given in Fig. 4. The log scale is taken to plot the MSEs for an easy identification of curves. On the whole, a good fit between the difference equations and the simulation can be seen in Fig. 4. For the strongest frequency component, the sinusoidal fluctuations can be identified with ease in both the tracking errors and MSEs. On the contrary, for the weakest sinusoid, vague and weak sinusoidal fluctuations can be found in the tracking errors, but almost none can be found in the tracking MSEs. This is because the portion in the tracking MSEs due to the additive noise is dominant and much larger than that generated by the FM.

B. Stability Bounds

Second, to obtain a stability bound for the LMS algorithm one needs to select one of the step size parameters as the grid search variable, and fix all the other step size values in the difference equations (16), (17) and (27), (28). Here, we consider a case with two frequency components. The step size \( \mu_1 \) is fixed while \( \mu_2 \) is regarded as the grid search variable whose stability bound is explored. The “step size” used in the grid search is 0.001. A comparison between the theoretical bounds and the simulated ones is given in Fig. 5, where an excellent agreement can be noticed. We changed the FM and the power of the frequency components, but the same results were obtained. This implies that the stability bounds are not affected by the FM as well as the signal power.

C. Steady-State MSEs

Third, the theoretical steady-state MSEs (42) are compared with their simulated values versus two step size parameters in Fig. 6. The theoretical and simulated MSEs are given in Fig. 6(a), (b), respectively, with the same axis ranges. Their differences are shown in Fig. 6(c). On the whole, theory presents very good agreement with the simulated values. The optimum
step size parameters by numerical grid search based on (42), theoretical steepest-descent recursive search based on the MSE minimization, and the simulation are summarized in Table I for comparison. The simulated step size values are in very good agreement with their theoretical counterparts. The update convergence of the two step size parameters are given in Fig. 7. Obviously, the step size initial values obtained from (49) are quite close to the optimum ones where the MSE shows the minimum.

V. CONCLUSION

The main results obtained in this work are as follows.

1) The estimation problem in the conventional LMS analyzer eventually becomes a tracking one due to the existence of the FM. This causes the performance degeneration of the LMS algorithm.

2) A set of approximately linear difference equations governing the dynamics of the algorithm in the presence of FM has been derived, which describe the convergence properties of the algorithm very well. Furthermore, upper bounds for a step size parameter can be achieved by fixing all the other step size parameters and performing grid search by using of the obtained difference equations.

3) The optimum step size parameters derived through the minimization of the steady-state tracking MSE reveal that larger step size values have to be used for larger FM and/or signal powers, while smaller ones are needed when the power of the additive noise is larger.

4) The FM also produces additional bias terms in the squared amplitude and phase estimates, which will significantly decrease the performance of the LMS when used as a detector.

<table>
<thead>
<tr>
<th>initial value by Eq.(49)</th>
<th>grid search by Eq.(42)</th>
<th>numerical (Fig.7)</th>
<th>simulated (Fig.6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_1 )</td>
<td>0.1165</td>
<td>0.1050</td>
<td>0.1075</td>
</tr>
<tr>
<td>( \mu_2 )</td>
<td>0.1849</td>
<td>0.1700</td>
<td>0.1704</td>
</tr>
</tbody>
</table>

Fig. 7. Recursive update of the step size parameters \( \mu_1 \) and \( \mu_2 \) that minimize the total MSE of (42), with the other parameters the same as in Fig. 6.

ACKNOWLEDGMENT

The authors would like to thank the five anonymous reviewers for their constructive comments and for bringing to the authors’ attention, some related references; and Miss A. Sudo, Fujitsu Systems Ltd., Tokyo, Japan, for her assistance in preparing part of the analytical equations.

REFERENCES


Yegui Xiao (M’96) received the B.S. degree in electrical engineering from Northeastern University, Shenyang, China, and the M.S. and Ph.D. degrees in electrical engineering from Hiroshima University, Hiroshima-Hiroshima, Japan, in 1984, 1988, and 1991, respectively.

After working with Fuji Electric Ltd., Tokyo, as a Senior Systems Engineer for two years, he joined Toyohashi University of Technology, as an Assistant Professor in 1993. He joined Saga University, Saga, Japan, as an Associate Professor in 1997. In 1999, he became an Associator Professor at Hiroshima Prefectural Women’s University (HPWU), Hiroshima. He has been a Professor at HPWU since April, 2004. He was a Visiting Professor in the Institute for Computing, Information and Cognitive Systems (ICICS) at the University of British Columbia, Vancouver, Canada, in 2002. His main research interests are in statistical signal processing, adaptive filters and their applications, image processing for face and facial expression recognition.

Dr. Xiao received the Best Thesis Work Award for his B.S. degree from the Department of Automatic Control, Northeastern University, Shenyang, China, and the Excellent Paper Presentation Award from the Institute of Electrical Engineers of Japan, in 1984 and 1991, respectively. He is a member of the Institute of Electronics, Information, and Communications Engineers (IEICE), Japan, the Society of Instrument and Control Engineers (SICE), Japan, and Information Processing Society of Japan.

Liying Ma received the B.S. degree from Northeastern University, Shenyang, China, in 1984, the M.S. degree from Hiroshima University, Higashi-Hiroshima, Japan, and the Ph.D. degree from Concordia University, Montreal, QC, Canada, in 1991 and 2001, respectively, both in electrical and computer engineering.

She was a Research Electrical Engineer with the Automation Research Institute, Ministry of Metallurgical Industry, Beijing, China, from 1984 to 1998. From 1991 to 1992, she worked as a Systems Engineer at Yokogawa Engineering Service Company, Tokyo, Japan. In 2002, she was a Scientific Engineer at the University of British Columbia, Vancouver, BC, Canada. She has been a Post-Doctoral Fellow at Concordia University since 2003. Her main areas of research interests are in neural networks and their applications to image processing and pattern recognition, bioinformatics, sensor and networked control systems, optimization and intelligent systems, and digital signal processing and applications.

Li Xu (M’92) received the B.Eng. degree in automatic control from Huazhong University of Science and Technology, Wuhan, China, in January 1982, and the M.Eng. degree in information and computer sciences and the Dr.Eng. degree in System and Information Engineering from Toyohashi University of Technology, Toyohashi, Japan, in March 1990 and March 1993, respectively.

From April 1993 to March 1998, he was an Assistant Professor in the Department of Knowledge-Based Information Engineering of Toyohashi University of Technology. From April 1998 to March 2000, he was a Lecturer at the Department of Information Management, School of Business Administration, Asahi University, Gifu, Japan. Since April 2000, he has been an Associate Professor at the Department of Electronics and Information Systems, Faculty of Systems Science and Technology, Akita Prefectural University, Akita, Japan. His research interests include multidimensional system theory, signal processing and the applications of computer algebra to system theory. He has been an Associate Editor for the journal Multidimensional Systems and Signal Processing (MSSP) since 2000, and was the Guest Co-editor for the special issue on “Applications of Grobner Bases in Multidimensional Systems and Signal Processing” for MSSP published in 2001.

Rabab Kreidieh Ward (S’71–M’72–SM’85–F’99) was born in Beirut, Lebanon. She received the B.Eng. degree from the University of Cairo, Cairo, Egypt, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1969 and 1972, respectively.

She is a Professor in the Electrical and Computer Engineering Department and the Director of the Institute for Computing, Information, and Cognitive Systems at the University of British Columbia, Vancouver, BC, Canada. Her expertise lies in digital signal processing and applications to cable TV, high definition TV, video compression and medical images, including mammography, microscopy and cell images. She holds six patents and many of her research ideas have been transferred to industry, including her invention of a non invasive testing system for cable TV, a nondestructive system to measure the size of fish for the aquaculture industry, and a compression scheme for photographic images. She has published over 200 papers and chapters in scientific books.

Dr. Ward was the general chair of ICIER 2000. She is a Fellow of the Royal Society of Canada, Fellow of the Engineering Institute of Canada, and Fellow of the Canadian Academy of Engineers. She is also a recipient of a UBC Killam Research Prize.

Akira Ikuta was born in Japan in November 1951. He received the B.E. and M.E. degrees from Kyushu Institute of Technology, Kyushu, Japan, in 1976 and 1978, respectively, and the D.Eng. degree from Hiroshima University, Higashi-Hiroshima, Japan, in 1981.

After working as a Research Associate of Electrical Engineering at the Hiroshima University and an Assistant Professor in Mechanical Systems Engineering at the Kinki University, he is now a Professor in Human Life and Environmental Science at the Hiroshima Prefectural Women’s University. His research interests include the digital signal/information processing and wave motion environment (sound, vibration and electromagnetic wave).

Dr. Ikuta is a member of the scientific committee of the International Institute of Acoustics and Vibration.