MEAN AND MEAN-SQUARE ANALYSIS OF THE COMPLEX LMS ALGORITHM FOR NON-CIRCULAR GAUSSIAN SIGNALS

Scott C. Douglas

Department of Electrical Engineering
Southern Methodist University
Dallas, Texas 75275 USA

Danilo P. Mandic

Department of Electrical and Electronic Engineering
Imperial College
London, SW7 2BT United Kingdom

ABSTRACT

The least-mean-square (LMS) algorithm is a useful and popular procedure for adaptive signal processing of both real-valued and complex-valued signals. Past analysis of the complex LMS algorithm has assumed that the input signal vector is circularly-distributed, such that the pseudo-covariance matrix of the input signal is zero. In this paper, we relax this assumption, providing a complete mean and mean-square analysis of the complex LMS algorithm for non-circular Gaussian signals. Our analysis unifies the statistical descriptions of the conventional (real-valued) LMS and complex LMS algorithms as specific cases of our more-general behavioral description, negating the need for a distinction between these two procedures. Simulations indicate that our analysis more-accurately predicts the behavior of complex LMS for non-circular signals as compared to existing analyses in the scientific literature.

1. INTRODUCTION

The least-mean-square (LMS) algorithm is perhaps the most-often-used adaptive signal processing algorithm. Based on a stochastic version of gradient descent applied to a mean-square error cost, the LMS algorithm is used in a number of applications in communications, image processing, speech processing, and medicine, among other disciplines. The algorithm is simple to code, computationally-efficient, and has robust performance. Numerous analyses of the performance of the LMS algorithm have been developed in the signal processing literature.

The complex LMS algorithm is an extension of the LMS algorithm to complex-valued signals [1, 2]. It is particularly useful in communication applications involving modulated signals, as such signals are best represented using complex numbers in a baseband representation. For the weight vector $\mathbf{w}(k) = [w_1(k) \cdots w_L(k)]^T$ at time k, the complex LMS algorithm is given by

$$y(k) = \mathbf{x}^H(k)\mathbf{w}(k) \quad (1)$$

$$e(k) = d(k) - y(k) \quad (2)$$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e(k)\mathbf{x}(k). \quad (3)$$

where $\mathbf{x}(k) = [x_1(k) \cdots x_L(k)]^T$ is the input signal vector, $y(k)$ is the output signal, $d(k)$ is the desired response signal, $e(k)$ is the error signal, and μ is the step size. In this form, the complex LMS algorithm differs from the LMS algorithm only in the use of the Hermitian transpose in the definition of $y(k)$.

The first analysis of the complex LMS algorithm was most-likely developed by K.D. Senne in his Ph.D. dissertation [3], which was later expanded on and published as an oft-cited paper [4]. A related analysis as applied to the adaptive line enhancer appears in [5].

The analysis in [4] assumes that the pair $\{d(k), \mathbf{x}(k)\}$ is jointly-Gaussian, and that $\mathbf{x}(k)$ is zero mean and complex circular Gaussian (that is, having a rotationally-invariant p.d.f. in the complex domain), such that

$$\mathbf{R} = E\{\mathbf{x}(k)\mathbf{x}^H(k)\} \neq \mathbf{0} \quad (4)$$

$$\text{and} \quad \mathbf{P} = E\{\mathbf{x}(k)\mathbf{x}^T(k)\} \equiv \mathbf{0}. \quad (5)$$

The matrices $\mathbf{R}$ and $\mathbf{P}$ are known as the covariance and pseudo-covariance matrices of $\mathbf{x}(k)$, respectively, when $\mathbf{x}(k)$ is zero mean. While the assumption that $\mathbf{P} \equiv \mathbf{0}$ is sometimes reasonable, it is not true in general. For example, when a real-valued BPSK signal is transmitted over a dispersive channel, the received signal to an equalizer can exhibit non-circular statistical behavior. Recently, the use of non-circular statistics has been leveraged to develop procedures for carrier recovery [6], direction-of-arrival estimation [7], and signal separation [8]. It is not clear, however, how the statistical behavior of the complex LMS algorithm changes when the input signal exhibits non-circular statistics.

In this paper, we provide a complete statistical analysis of the complex LMS algorithm for complex non-circular jointly Gaussian input and desired response signals, such that $\mathbf{P} \neq \mathbf{0}$. This analysis includes the work of Horowitz and Senne in [4] as a special case when $\mathbf{P} = \mathbf{0}$. Interestingly, it also leads to a *unifying analysis* of the real-valued LMS and complex LMS algorithms, as it includes the well-known analysis by Feuer and Weinstein [9] in the case of real-valued signals such that $\mathbf{R} = \mathbf{P} = \mathbf{P}^*$. Thus, there is no need to present two different analyses for the real-valued and complex LMS algorithms, negating the main reason for a distinction between the real-valued and complex-valued data scenarios. From our study, several issues are illuminated about the complex LMS algorithm:

- Its mean behavior is unaffected by input signal non-circularity and desired-response signal non-circularity.
- Its mean-square behavior is independent of the non-circularity of $d(k)$.
- Its mean-square behavior depends weakly on the non-circularity of the input signal vector. This effect is most-pronounced for small filter lengths L and highly-non-circular input signals.
- The modes of convergence of the algorithm cannot be decoupled, except for special cases of input signal non-circularity.

Simulations show that our analysis is accurate in predicting mean-square performance.

978-1-4244-3677-4/09/$25.00 ©2009 IEEE

2. SIMPLIFYING AND STRUCTURAL ASSUMPTIONS

For our analysis, we can assume without loss of generality that $d(k)$ is generated from the linear model

$$d(k) = \mathbf{x}^H(k)\mathbf{w}_{opt} + \eta(k), \quad (6)$$

where $\eta(k)$ is complex non-circular Gaussian and uncorrelated with $\mathbf{x}(k)$, because of the joint Gaussianity of $d(k)$ and $\mathbf{x}(k)$.

Throughout our analysis, we will make use of the independence assumptions, such that $\{d(l), \mathbf{x}(l)\}$ is independent of $\{d(k), \mathbf{x}(k)\}$ when $k \neq l$. While not strictly true in some cases such as FIR filtering, such assumptions lead to reasonably-accurate behavior, especially for small step sizes μ.

3. MEAN BEHAVIOR

Define the coefficient error vector

$$\mathbf{v}(k) = \mathbf{w}(k) - \mathbf{w}_{opt}. \quad (7)$$

where $\mathbf{w}_{opt}$ is We can then write the update for $\mathbf{w}(k)$ in terms of $\mathbf{v}(k)$ as

$$\mathbf{v}(k+1) = \mathbf{v}(k) - \mu\mathbf{x}(k)\mathbf{x}^H(k)\mathbf{v}(k) + \mu\eta(k)\mathbf{x}(k). \quad (8)$$

Taking expectations of both sides of (8) and employing the independence assumptions, we obtain

$$E\{\mathbf{v}(k+1)\} = (\mathbf{I} - \mu\mathbf{R})E\{\mathbf{v}(k)\}, \quad (9)$$

which is identical to that for complex LMS with circularly-symmetric Gaussian input signals. Hence, the mean characteristics of complex LMS do not change with non-circular input signals.

4. MEAN-SQUARE BEHAVIOR

We now develop analytical equations for the evolution of the coefficient error correlation matrix

$$\mathbf{K}(k) = E\{\mathbf{v}(k)\mathbf{v}^H(k)\}. \quad (10)$$

This matrix can be used to determine the MSE for any k through the relations

$$E\{|e(k)|^2\} = \sigma_\eta^2 + \xi_{EMSE}(k) \quad (11)$$
$$\xi_{EMSE}(k) = \mathrm{tr}[\mathbf{R}\mathbf{K}(k)] \quad (12)$$

where $\xi_{EMSE}(k)$ is the excess MSE at time k. Taking both sides of (8) and multiplying them by their Hermitian transposes, we obtain

$$\begin{aligned}\mathbf{v}(k+1)\mathbf{v}^H(k+1) &= \mathbf{v}(k)\mathbf{v}^H(k) - \mu\mathbf{x}(k)\mathbf{x}^H(k)\mathbf{v}(k)\mathbf{v}^H(k) \\ &\quad - \mu\mathbf{v}(k)\mathbf{v}^H(k)\mathbf{x}(k)\mathbf{x}^H(k) + \mu^2|\eta(k)|^2\mathbf{x}(k)\mathbf{x}^H(k) \\ &\quad + \mu^2\mathbf{x}(k)\mathbf{x}^H(k)\mathbf{v}(k)\mathbf{v}^H(k)\mathbf{x}(k)\mathbf{x}^H(k) \\ &\quad + (\text{crossterms}), \end{aligned} \quad (13)$$

where the crossterms not shown in the above relation are zero under expectation with the independence assumptions in place.

Taking expectations of both sides of the above relation, we get

$$\begin{aligned}\mathbf{K}(k+1) &= \mathbf{K}(k) - \mu[\mathbf{R}\mathbf{K}(k) + \mathbf{K}(k)\mathbf{R}] + \mu^2\sigma_\eta^2\mathbf{R} \\ &\quad + \mu^2 E\{\mathbf{x}(k)\mathbf{x}^H(k)\mathbf{v}(k)\mathbf{v}^H(k)\mathbf{x}(k)\mathbf{x}^H(k)\}\end{aligned} \quad (14)$$

We look at the (i,j)th entry of the expectation matrix of the last term in (14) under the independence assumptions.

$$\begin{aligned}&\left[E\{\mathbf{x}(k)\mathbf{x}^H(k)\mathbf{v}(k)\mathbf{v}^H(k)\mathbf{x}(k)\mathbf{x}^H(k)\}\right]_{ij} \\ &= \sum_{l=1}^{L}\sum_{m=1}^{L} E\{x_i(k)x_l^*(k)x_m(k)x_j^*(k)\}E\{v_l(k)v_m^*(k)\}\end{aligned} \quad (15)$$

The fourth-moment term has the value

$$E\{x_i(k)x_l^*(k)x_m(k)x_j^*(k)\} = r_{il}r_{mj} + p_{im}p_{lj}^* + r_{ij}r_{ml} \quad (16)$$

using the fourth-moment expression for jointly Gaussian random variables. Thus, we have

$$\begin{aligned}&E\{\mathbf{x}(k)\mathbf{x}^H(k)\mathbf{v}(k)\mathbf{v}^H(k)\mathbf{x}(k)\mathbf{x}^H(k)\} \\ &= \mathbf{R}\mathbf{K}(k)\mathbf{R} + \mathbf{P}\mathbf{K}^*(k)\mathbf{P}^* + \mathbf{R}\mathrm{tr}[\mathbf{R}\mathbf{K}(k)]\end{aligned} \quad (17)$$

where tr[·] denotes the trace of a matrix. Thus, the evolution of the coefficient error correlation matrix is given by the equation

$$\begin{aligned}\mathbf{K}(k+1) &= \mathbf{K}(k) - \mu[\mathbf{R}\mathbf{K}(k) + \mathbf{K}(k)\mathbf{R}] \\ &\quad + \mu^2(\sigma_\eta^2 + \mathrm{tr}[\mathbf{R}\mathbf{K}(k)])\mathbf{R} \\ &\quad + \mu^2\mathbf{R}\mathbf{K}(k)\mathbf{R} + \mu^2\mathbf{P}\mathbf{K}^*(k)\mathbf{P}^*.\end{aligned} \quad (18)$$

Several remarks can be made at this point:

1. The update relation in (18) unites the complex and real-valued input signal case. If $\mathbf{x}(k)$ is completely real-valued, then $\mathbf{R} = \mathbf{P} = \mathbf{P}^*$, and we obtain the well-known relations originally published by A. Feuer and E. Weinstein [9]. If $\mathbf{x}(k)$ is complex circular, then $\mathbf{P} = \mathbf{0}$, such that we obtain the well-known relations originally published by Horowitz and Senne [4]. It also includes cases that neither analysis can examine: those involving non-circular, non-real-valued input signals.

2. The expression in (18) is not diagonalizable in general. If we define the eigenvalue decomposition

$$\mathbf{R} = \mathbf{Q}\mathbf{\Sigma}^2\mathbf{Q}^H \quad (19)$$

where $\mathbf{Q}$ is orthonormal and $\mathbf{\Sigma}^2$ is a diagonal matrix of eigenvalues, it is not the case that $\mathbf{Q}^H\mathbf{P}\mathbf{Q}^*$ is diagonal in general. Thus, it is not possible to diagonalize (18) in an attempt to decouple the evolutions of the diagonal elements of $\mathbf{Q}^H\mathbf{K}(k)\mathbf{Q}$ as is done for both the real-valued and the circular complex data cases.

3. If μ is small, then we can neglect the terms that depend on $\mu^2\mathbf{K}(k)$. Thus, we obtain relations that are identical to the small step size case for the standard analysis of the complex LMS algorithm. Hence, the excess MSE of the complex LMS algorithm is unaffected by the non-circularity of the data to first order (*i.e.* for small step sizes).

5. SPECIAL CASES OF INTEREST

In order to understand the convergence behavior of (18), it is useful to assume that $\mathbf{x}(k)$ possesses additional statistical structure. We now describe two such specific cases of interest.

5.1. Case A: Uncorrelated Non-Circular Input Data

Assume that

$$\mathbf{R} = \sigma^2\mathbf{I} \tag{20}$$

such that the input signal vector is circularly uncorrelated. Then, it can be shown that it is always possible to select a complex orthonormal transformation matrix $\mathbf{Q}$ such that

$$\mathbf{Q}^H\mathbf{P}\mathbf{Q}^* = \sigma^2\mathbf{\Lambda}, \tag{21}$$

where $\mathbf{\Lambda}$ is a diagonal matrix of noncircularity coefficients that satisfy $0 \leq \lambda \leq 1$. Such a transformation is a special case of the strong uncorrelating transform [8]. Define

$$\widetilde{\mathbf{K}}(k) = \mathbf{Q}^H\mathbf{K}(k)\mathbf{Q}. \tag{22}$$

Then, we can transform (18) to

$$\begin{aligned}\widetilde{\mathbf{K}}(k+1) &= (1-\mu\sigma^2)^2\widetilde{\mathbf{K}}(k) + \mu^2\sigma^2(\sigma_\eta^2 + \sigma^2\mathrm{tr}[\widetilde{\mathbf{K}}(k)])\mathbf{I} \\ &\quad + \mu^2\mathbf{\Lambda}\widetilde{\mathbf{K}}^*(k)\mathbf{\Lambda}. \end{aligned}\tag{23}$$

The off-diagonal elements of $\widetilde{\mathbf{K}}(k)$ are decoupled and evolve as

$$\kappa_{ij}(k+1) = (1-\mu\sigma^2)^2\kappa_{ij}(k) + \mu^2\sigma^4\lambda_i\lambda_j\kappa_{ij}^*(k). \tag{24}$$

Because they have no driving term, then $\lim_{k\to\infty}\kappa_{ij}(k) = 0$ for small enough values of μ. The on-diagonal terms are real-valued and evolve according to a vector relation. Let

$$\mathbf{s}(k) = [\kappa_{11}(k)\cdots\kappa_{LL}(k)]^T. \tag{25}$$

Then, we have

$$\begin{aligned}\mathbf{s}(k+1) &= \left[(1-\mu\sigma^2)^2\mathbf{I} + \mu^2\sigma^4\mathbf{\Lambda}^2 + \mu^2\sigma^4\mathbf{1}\mathbf{1}^T\right]\mathbf{s}(k) \\ &\quad + \mu^2\sigma^2\sigma_\eta^2\mathbf{1}. \end{aligned}\tag{26}$$

The steady-state value of $\mathbf{s}(k)$ is given by

$$\mathbf{s}_{ss} = \mu\sigma_\eta^2[2\mathbf{I} - \mu\sigma^2\left\{\mathbf{I} + \mathbf{\Lambda}^2 + \mathbf{1}\mathbf{1}^T\right\}]^{-1}\mathbf{1}. \tag{27}$$

Employing the matrix inversion lemma, we obtain

$$s_{ss,i} = \frac{\dfrac{\mu\sigma_\eta^2}{2-\mu\sigma^2(1+\lambda_i^2)}}{1-\displaystyle\sum_{j=1}^{L}\frac{\mu\sigma^2}{2-\mu\sigma^2(1+\lambda_j^2)}} \tag{28}$$

The steady-state excess MSE is

$$\xi_{EMSE,ss} = \sigma^2\mathbf{s}_{ss}^T\mathbf{1} = \frac{\displaystyle\sum_{i=1}^{L}\frac{\mu\sigma^2\sigma_\eta^2}{2-\mu\sigma^2(1+\lambda_i^2)}}{1-\displaystyle\sum_{j=1}^{L}\frac{\mu\sigma^2}{2-\mu\sigma^2(1+\lambda_j^2)}}, \tag{29}$$

and the misadjustment is

$$M = \frac{\xi_{EMSE,ss}}{\sigma_\eta^2} = \frac{\displaystyle\sum_{i=1}^{L}\frac{\mu\sigma^2}{2-\mu\sigma^2(1+\lambda_i^2)}}{1-\displaystyle\sum_{j=1}^{L}\frac{\mu\sigma^2}{2-\mu\sigma^2(1+\lambda_j^2)}} \tag{30}$$

Consider the additional simplification that all of the circularity coefficients are the same, such that $\lambda_i = \lambda$, $1 \leq i \leq L$. Then, we have

$$M = \frac{\mu\sigma^2 L}{2-\mu\sigma^2[L+1+\lambda^2]}. \tag{31}$$

Again, this result unifies the complex and real-valued data cases, as $\lambda = 0$ yields the complex-valued result for misadjustment and $\lambda = 1$ yields the corresponding real-valued result.

5.2. Case B: A Specific Correlated Non-Circular Input Data Model

Suppose that the input signal vector has a uniform non-circularity coefficient λ in all of its dimensions, such that

$$\mathbf{R} = \mathbf{C}\mathbf{C}^H \tag{32}$$

$$\mathbf{P} = \lambda\mathbf{C}\mathbf{C}^T \tag{33}$$

where $\mathbf{C}$ is any strong uncorrelating transform for the data. Then, using the singular value decomposition of $\mathbf{C}$, it can be shown that

$$\mathbf{R} = \mathbf{Q}\mathbf{\Sigma}^2\mathbf{Q}^H \tag{34}$$

$$\mathbf{P} = \lambda\mathbf{Q}\mathbf{\Sigma}^2\mathbf{Q}^T \tag{35}$$

where $\mathbf{Q}$ contains the left singular vectors of $\mathbf{C}$ and $\mathbf{\Sigma}$ contains its corresponding singular values. Then, we may again transform $\mathbf{K}(k)$ to $\widetilde{\mathbf{K}}(k)$ to obtain the evolutionary relation

$$\begin{aligned}\widetilde{\mathbf{K}}(k+1) &= \widetilde{\mathbf{K}}(k) - \mu\left[\mathbf{\Sigma}^2\widetilde{\mathbf{K}}(k) + \widetilde{\mathbf{K}}(k)\mathbf{\Sigma}^2\right] \\ &\quad + \mu^2\left[\sigma_\eta^2 + \mathrm{tr}[\mathbf{\Sigma}^2\widetilde{\mathbf{K}}(k)]\right]\mathbf{\Sigma}^2 \\ &\quad + \mu^2\mathbf{\Sigma}^2\left[\widetilde{\mathbf{K}}(k) + \lambda^2\widetilde{\mathbf{K}}^*(k)\right]\mathbf{\Sigma}^2 \end{aligned}\tag{36}$$

The off-diagonal terms are decoupled from each other and the on-diagonal terms, and they evolve as

$$\kappa_{ij}(k+1) = (1-\mu\sigma_i^2)\kappa_{ij}(k)(1-\mu\sigma_j^2) + \mu^2\lambda^2\sigma_i^2\sigma_j^2\kappa_{ij}^*(k) \tag{37}$$

As before, if the step size μ is small enough, $\kappa_{ij}(k)$ decays exponentially to zero for $i \neq j$. For the on-diagonal terms,

$$\mathbf{s}(k+1) = \left[(\mathbf{I}-\mu\mathbf{\Sigma}^2)^2 + \mu^2\lambda^2\mathbf{\Sigma}^4 + \mu^2\mathbf{p}\mathbf{p}^T\right]\mathbf{s}(k) + \mu^2\sigma_\eta^2\mathbf{p}, \tag{38}$$

where

$$\mathbf{p} = [\sigma_1^2\cdots\sigma_L^2]^T \tag{39}$$

The steady-state value of $\mathbf{s}(k)$ is

$$\mathbf{s}_{ss} = \mu\sigma_\eta^2\left[2\mathbf{\Sigma}^2 - \mu(1+\lambda^2)\mathbf{\Sigma}^4 - \mu\mathbf{p}\mathbf{p}^T\right]^{-1}\mathbf{p}. \tag{40}$$

Employing the matrix inversion lemma, we obtain

$$s_{ss,i} = \frac{\dfrac{\mu\sigma_\eta^2}{2-\mu\sigma_i^2(1+\lambda^2)}}{1-\displaystyle\sum_{j=1}^{L}\frac{\mu\sigma_j^2}{2-\mu\sigma_j^2(1+\lambda^2)}} \tag{41}$$

The steady-state excess MSE is

$$\xi_{EMSE,ss} = \mathbf{s}_{ss}^T \mathbf{p} = \frac{\sum_{i=1}^{L} \frac{\mu\sigma_i^2\sigma_\eta^2}{2-\mu\sigma_i^2(1+\lambda^2)}}{1-\sum_{j=1}^{L} \frac{\mu\sigma_j^2}{2-\mu\sigma_j^2(1+\lambda^2)}}, \quad (42)$$

and the misadjustment is

$$M = \frac{\xi_{EMSE,ss}}{\sigma_\eta^2} = \frac{\sum_{i=1}^{L} \frac{\mu\sigma_i^2}{2-\mu\sigma_i^2(1+\lambda^2)}}{1-\sum_{j=1}^{L} \frac{\mu\sigma_j^2}{2-\mu\sigma_j^2(1+\lambda^2)}} \quad (43)$$

Again, the expressions derived in this case bridge the gap between the real-valued ($\lambda = 1$) and circular complex-valued ($\lambda = 0$) data cases.

6. NUMERICAL EVALUATIONS

We now explore the accuracy of the above analysis through numerical evaluations. The scenarios considered in this section employ non-circular Gaussian distributions for both the input signal vector and desired response signal, respectively. We use the existing analysis for the complex LMS algorithm in [4] and compare its predictive accuracy to that of the analysis in this paper, although [4] does not allow for non-circular signal distributions within its statistical signal models. Thus, we can observe both the accuracy of our new results in predicting actual convergence behavior as well as the performance differences in LMS convergence behavior due to non-circular signals. We do not employ the analysis in [9] as it is not specified for complex-valued signals and filter coefficients. In each case, 100000 different simulation runs of the complex LMS algorithm were run and estimates of the quantities of interest were ensemble-averaged from the simulations to obtain the observed data with which the analyses can be compared.

Consider first a simple single-channel adaptive FIR filtering tasks, in which $L = 4$, $\mu = 0.15$,

$$\mathbf{w}_{opt} = [0.25 + j\ 0.5 + 0.75j\ 0.75 + 0.5j\ 1 + 0.25j]^T, \quad (44)$$

$x(k)$ is a zero-mean non-circular Gaussian signal with $E\{x(k)x^*(k-j)\} = \delta(j)$ and $E\{x(k)x(k-j)\} = 0.5\delta(j)$, and $\eta(k)$ is a zero-mean circular Gaussian signal with $E\{\eta(k)\eta^*(k-j)\} = 0.01\delta(j)$. In this case, the analysis in Case A of the last section is appropriate, where $\Lambda = 0.5\mathbf{I}$. In these simulations, $\mathbf{w}(0)$ was chosen as $\mathbf{w}_{opt} + \mathbf{v}(0)$, where $\mathbf{v}(0)$ is circular-complex Gaussian-distributed with zero mean and covariance $\mathbf{K}(0) = \mathbf{I}$. Shown in Figure 1 is the evolution of the average normalized coefficient error power $E\{\text{tr}[\mathbf{K}(k)]\}/E\{\text{tr}[\mathbf{K}(0)]\}$ using Eq. (18) from this paper, along with that generated from the analysis in [4]. As can be seen, our analysis is the most accurate in predicting the simulated behavior because ours takes into account the non-circularity of the input and desired response signals. The main performance difference is a lower steady-state value for $E\{\mathbf{K}(k)\}$. Considering the average of the last 20 points of the averaged simulation runs to be the "true value" of this quantity, our analysis accurately predicts this value to within -0.00073dB. The analysis in [4] predicts a value that differs by -0.324 dB in this case.

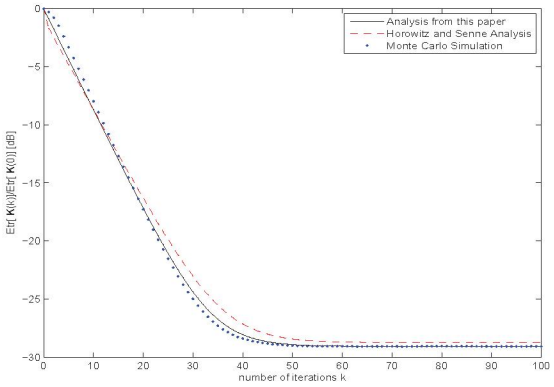

Fig. 1. Evolutions of the normalized total coefficient error power from simulation and from two different analytical predictions of performance for the single-channel FIR filtering example.

We now consider an adaptive array processing task, in which $\mathbf{x}(k)$ is generated according to the critically-spaced uniform linear array (ULA) model with three users given by

$$\mathbf{x}(k) = \mathbf{u}(k) + \sum_{p=1}^{3} \mathbf{v}(\phi_p) 10^{SNR_p/20} s_i(k), \quad (45)$$

where $\mathbf{v}(\phi_p)$ is the steering vector for the pth user whose ith element is $\sqrt{1/L}\exp(-j\pi(i-1)\sin(\phi_p))$, $s_p(k)$ is the message signal from the pth user with unit variance, SNR_p is the signal-to-noise ratio of the pth user, and $\mathbf{u}(k)$ is the uncorrelated circular Gaussian measurement noise vector at time k with $E\{\mathbf{u}(k)\mathbf{u}^T(k)\} = \sigma_u^2\mathbf{I}$ and $\sigma_u^2 = 1$. For these simulations, $s_1(k)$ is generated as a circular-distributed complex Gaussian source, whereas $s_2(k)$ and $s_3(k)$ are real-valued Gaussian sources. Moreover, we choose the directions-of-arrival $[\phi_1\ \phi_2\ \phi_3] = [-10^\circ\ \ 10^\circ\ \ 25^\circ]$ and the signal-to-noise ratios $[SNR_1\ SNR_2\ SNR_3] = [10\ 20\ 25]$ dB, respectively. Thus, the covariance and pseudo-covariance matrices are

$$\mathbf{R} = \mathbf{I} + \sum_{p=1}^{3} 10^{SNR_p/10} \mathbf{v}(\phi_p)\mathbf{v}^H(\phi_p) \quad (46)$$

$$\mathbf{P} = \sum_{p=2}^{3} 10^{SNR_p/10} \mathbf{v}(\phi_p)\mathbf{v}^T(\phi_p), \quad (47)$$

respectively. For the complex LMS algorithm, we set $d(k) = s_1^*(k)$, such that the statistical model used in this paper is accurate, with

$$\sigma_\eta^2 = 1 - \mathbf{w}_{opt}^H \mathbf{R} \mathbf{w}_{opt} \quad (48)$$

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}[10^{SNR_2/20}\mathbf{v}(\phi_1)] \quad (49)$$

Shown in Figure 2 is the convergence of $E\{\text{tr}[\mathbf{K}(k)]\}/E\{\text{tr}[\mathbf{K}(0)]\}$ using Eq. (18) from this paper, along with that generated from the analysis in [4]. As can be seen, our analysis more accurately predicts overall convergence performance than that of [4]. The difference between what our analysis predicts for the final value in the plots and the estimated final value from the last 100 simulation points is -0.222dB, whereas the difference computed using the analysis in [4] is -0.880dB.

Given that we have an accurate analysis of the LMS algorithm for non-circular complex signals, one question that naturally arises

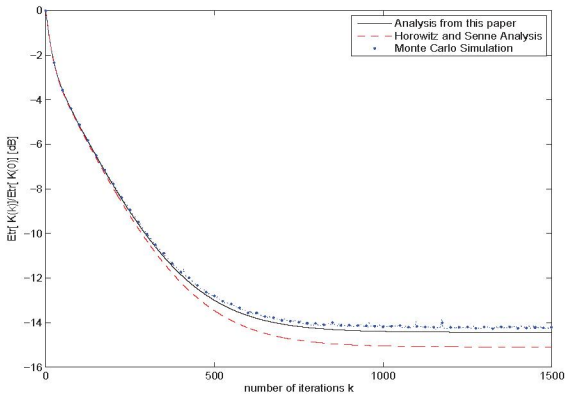

Fig. 2. Evolutions of the normalized total coefficient error power from simulation and from two different analytical predictions of performance for the multichannel array processing example.

is: How is the overall performance of the LMS adaptive filter affected by input signal noncircularity? This question could be explored in many ways given our analytical results. In order to simplify the discussion, consider the simplest case in which the input signal is i.i.d. with a constant circularity coefficient λ, such that the results of Section 5.1 are appropriate and $\mathbf{P} = \lambda\sigma^2\mathbf{I}$.

Assume that the step size of the LMS algorithm has been designed according to the circular data case, such that $\lambda = 0$ has been assumed. In this particular scenario, it is possible to determine both the increase in convergence speed and the increase in misadjustment due to a non-zero λ value. Figures 3 and 4 show the percentage increases in convergence time and final misadjustment in steady state for an $L = 2$-tap adaptive filter as a function of the step size ratio μ/μ_{opt} for various values of λ, where

$$\mu_{opt} = \frac{1}{L+1} \tag{50}$$

is the step size that yields the fastest convegence speed for circular i.i.d. input signals in this scenario. As can be seen, a non-circular input signal can lead to a 60% increase in convergence time as well as a 50% increase in final misadjustment as compared to the circular case in this situation, and the performance decreases are most severe for large step size values and non-circularity values. These effects decrease significantly, however, as the filter length is increased. Figures 5 and 6 show the percentage increases in convergence time and final misadjustment in steady state for an $L = 10$-tap adaptive filter, in which the convegence time and final misadjustment increase by no more than 10.5% and 10%, respectively. Figures 7 and 8 show the percentage increases in convergence time and final misadjustment in steady state for an $L = 50$-tap adaptive filter, in which the convegence time and final misadjustment increase by no more than 2% in practice. These results indicate that the effect of input signal noncircularity on overall convergence performance is relatively minor, particularly for long filter lengths L.

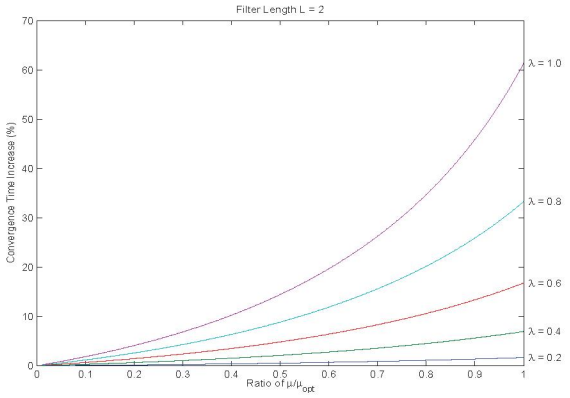

Fig. 3. Percentage increase in convergence time as a function of step size and input signal noncircularity, i.i.d. input signals, $L = 2$.

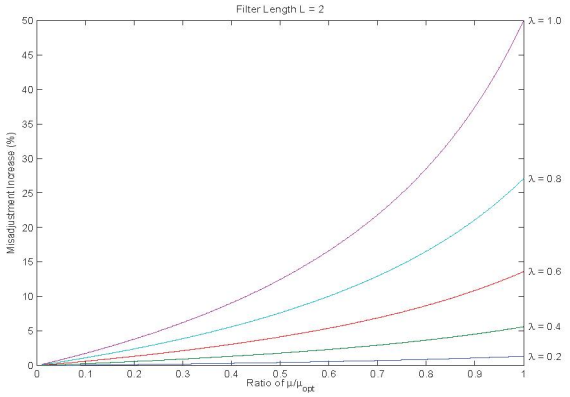

Fig. 4. Percentage increase in final misadjustment as a function of step size and input signal noncircularity, i.i.d. input signals, $L = 2$.

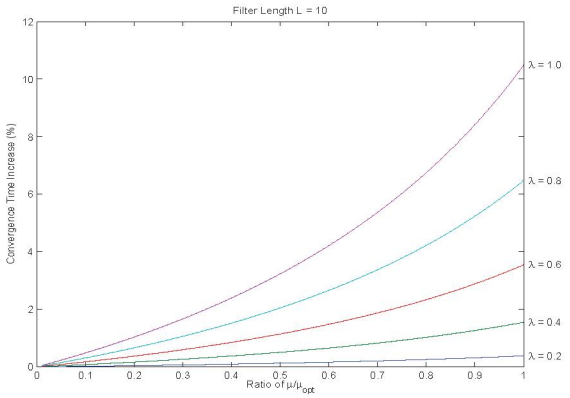

Fig. 5. Percentage increase in convergence time as a function of step size and input signal noncircularity, i.i.d. input signals, $L = 10$.

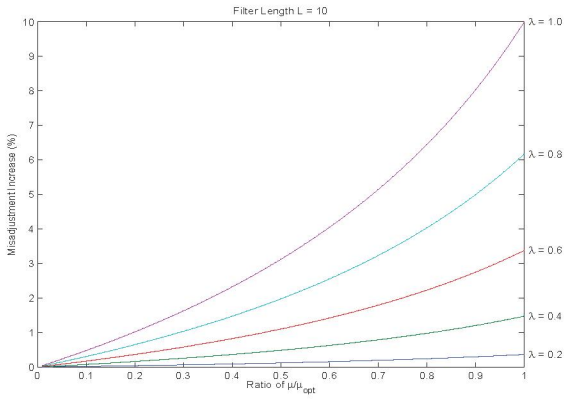

Fig. 6. Percentage increase in final misadjustment as a function of step size and input signal noncircularity, i.i.d. input signals, $L = 10$.

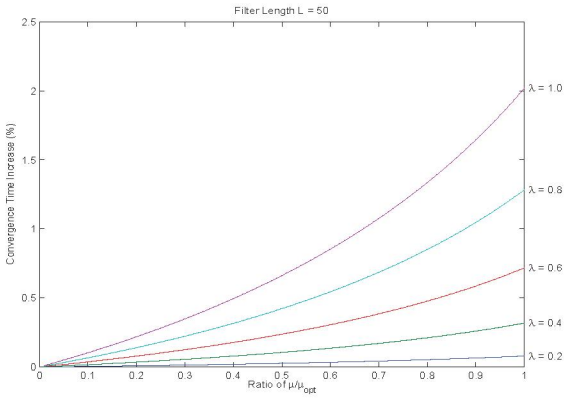

Fig. 7. Percentage increase in convergence time as a function of step size and input signal noncircularity, i.i.d. input signals, $L = 50$.

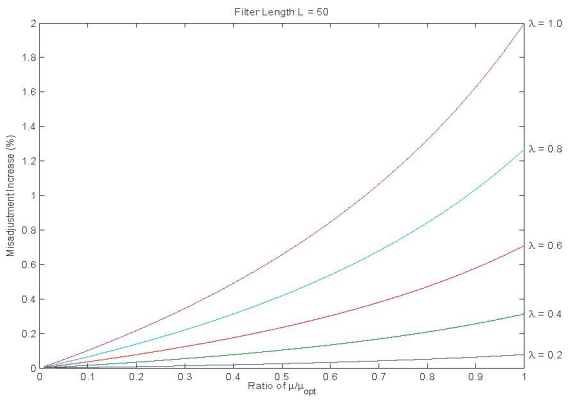

Fig. 8. Percentage increase in final misadjustment as a function of step size and input signal noncircularity, i.i.d. input signals, $L = 50$.

7. CONCLUSIONS

Previously, the least-mean-square (LMS) algorithm has been considered to be different when applied to real-valued and complex-valued data, with different derivations, performance analyses, and behavioral descriptions. In this paper, we provide an analysis for the LMS algorithm assuming complex non-circular Gaussian input and desired response signals. This analysis includes as special cases those published in [9] and [4]. Our analysis shows that the mean-square behavior of the modes of convergence of the complex LMS algorithm cannot be as easily decoupled as in other data cases, although specific statistical structures for the covariance and pseudo-covariance input signal matrix allow similar expressions for final misadjustment and excess mean-squared error to be derived. Simulations show that our new analysis more-accurately predicts convergence behavior when non-circular input and desired response signals are applied. In addition, these performance differences are most significant for large step size values and small filter lengths.

8. REFERENCES

[1] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proc. IEEE,* vol. 63, pp. 719-720, Apr. 1975.

[2] N.J. Bershad, "On the real and complex least mean square adaptive filter algorithms," *Proc. IEEE,* vol. 69, pp. 469-470, Apr. 1981.

[3] K.D. Senne, "Adaptive linear discrete time estimation," Ph.D. thesis, Stanford University, Stanford, CA, June 1968.

[4] L. Horowitz and K. Senne, "Performance advantage of complex LMS for controlling narrow-band adaptive arrays," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 29, pp. 722-736, June 1982.

[5] B. Fisher and N. Bershad, "The complex LMS adaptive algorithm–Transient weight mean and covariance with applications to the ALE," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 31, pp. 34-44, Feb. 1983.

[6] P. Ciblat and L. Vandendorpe, '1Blind carrier frequency offset estimation for noncircular constellation-based transmissions," *IEEE Trans. Signal Processing,* vol. 51, pp. 1378-1389, May 2003.

[7] J.-P. Delmas and H. Abeida, "Stochastic Cramer-Rao bound for noncircular signals with application to DOA estimation," *IEEE Trans. Signal Processing,* vol. 52, pp. 3192-3199, Nov. 2004.

[8] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: identifiability, uniqueness, and separability," *IEEE Trans. Inform. Theory,* vol. 52, pp. 1017-1029, Mar. 2006.

[9] A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 33, pp. 222-230, Feb. 1985.