THE LEAST-MEAN-MAGNITUDE-PHASE ALGORITHM WITH APPLICATIONS TO COMMUNICATIONS SYSTEMS

Scott C. Douglas[1,2]

[1]LGT Corporation
[2]Southern Methodist University
Dallas, Texas 75275 USA
douglas@lyle.smu.edu

Danilo P. Mandic

Imperial College London
Department of Electrical Engineering
London, SW7 2BT United Kingdom
d.mandic@imperial.ac.uk

ABSTRACT

This paper presents the least-mean-magnitude-phase (LMMP) algorithm for adaptive signal processing of complex-valued signals. The algorithm employs a decomposition of the mean-squared error cost and allows different normalized step sizes to be selected for the magnitude and phase errors. Simulations show that the algorithm is useful when either the amplitude or the phase relationship between the input and desired response signals can be accurately estimated.

Index Terms— Adaptive algorithm, adaptive equalizers, adaptive signal processing, adaptive systems, antenna arrays

1. INTRODUCTION

The least-mean-square (LMS) algorithm is perhaps the most popular adaptive filtering algorithm in existence. Its computational simplicity, ease of use, and consistency of behavior make it an excellent first choice in numerous signal processing applications. Originally specified for real-valued signals, it was extended to complex-valued signals in [1]. The complex LMS (CLMS) algorithm has proven to be particularly useful in communications and array processing.

There are a number of situations where the complex LMS algorithm does not perform adequately, particularly when the input and/or desired response signals exhibit unusual behavior in either their complex amplitudes or complex phases. One case is when the desired response signal undergoes a complex phase shift due to Doppler effects. To this end, Rupp proposed the constant-modulus channel estimator (CMCE) algorithm and showed its usefulness [2]. Another case is when the main information about the desired response signal is in its phase component. To this end, Tarighat and Sayed proposed the least-mean-phase (LMP) algorithm and indicated its usefulness [3].

In this paper, we propose an algorithm which we call the *least-mean-magnitude-phase (LMMP)* algorithm. The LMMP algorithm decomposes the instantaneous squared-error cost into the sum of two terms that are most-closely related to the amplitude-only error and the phase-only error, respectively, between the desired signal and the adaptive filter's output. The LMMP algorithm includes the CLMS algorithm as a special case, and a normalized version of the algorithm is easily specified. Unlike the CMCE algorithm, the LMMP algorithm is sensitive to phase. Unlike the LMP algorithm, the LMMP algorithm uses different step sizes to *directly* control the magnitude and phase error components of the MSE cost during adaptation. The LMMP algorithm exhibits excellent and robust convergence behavior when step size normalization is used. Most importantly, the LMMP algorithm can outperform the LMS algorithm in situations where portions of the input or desired response signals exhibit either amplitude or phase variations that are easily estimated or tracked. Simulation examples corresponding to these scenarios illustrate the new algorithm's advantages.

2. DERIVATION

Let $\mathbf{w}_k = [w_{1,k} \cdots w_{L,k}]^T$ denote the weight vector of a linear adaptive system with complex coefficients at time k. The complex-valued input signal vector $\mathbf{x}_k = [x_{1,k} \cdots x_{L,k}]^T$ is used to model a desired response signal d_k via the relation $y_k = \mathbf{w}_k^T \mathbf{x}_k$. For the moment, we neglect the time indices of these scalars and vectors in the following.

Consider the instantaneous least-mean-square cost:

$$J_{\mathrm{LMS}}(d, y) \quad = \quad |d - y|^2 \tag{1}$$

We seek a decomposition of this cost that splits it into the sum of two costs: one that is minimized when $|d| = |y|$, and one that is minimized when $\angle d = \angle y$. To this end, we expand the cost function in the complex domain to get

$$.J_{\mathrm{LMS}}(d, y) \quad = \quad |d|^2 + |y|^2 - 2|d||y|\cos(\angle d - \angle y), \tag{2}$$

where we have used the polarization identity for complex vectors to express $dy^* + d^* y$. We can now add and subtract $2|d||y|$ from the right-hand-side of (2) and group terms to get

$$\begin{aligned} J_{\mathrm{LMS}}(d, y) &= (|d| - |y|)^2 + 2|d||y|[1 - \cos(\angle d - \angle y)] && (3) \\ &= J_{\mathrm{m}}(d, y) + J_{\mathrm{p}}(d, y), && (4) \end{aligned}$$

where we have defined the magnitude and phase costs as

$$\begin{aligned} J_{\mathrm{m}}(d, y) &= (|d| - |y|)^2 && (5) \\ J_{\mathrm{p}}(d, y) &= 2|d||y|(1 - \cos(\angle d - \angle y)). && (6) \end{aligned}$$

Remark 1: The value of $J_{\mathrm{m}}(d,y)$ is zero when $|d| = |y|$, and the value of $J_{\mathrm{p}}(d,y)$ is zero when $\angle d = \angle y$. Thus, the goals of decomposing of $J_{\mathrm{LMS}}(d,y)$ have been met. The only caveat we have with $J_{\mathrm{p}}(d,y)$ is that it is also minimized when $|y| = 0$. Fortunately, the minimization of $J_{\mathrm{m}}(d,y)$ tends to force the value of $|y_k|$ away from the origin in practice, so long as $|d_k|$ does not remain at the complex origin for many consecutive samples.

Remark 2: Gradient minimization of $J_{\mathrm{m}}(d,y) + J_{\mathrm{p}}(d,y)$ yields the complex LMS algorithm. The key idea of this work is the design of *novel* criteria employing *different* linear combinations of $J_{\mathrm{m}}(d,y)$ and $J_{\mathrm{p}}(d,y)$ that lead to a new algorithm. To this end, consider the gradient of $J_{\mathrm{m}}(d,y)$ with respect to $\mathbf{w}^*$, given by is

$$\begin{aligned}\nabla_{\mathbf{w}^*} J_{\mathrm{m}}(d,y) &= -2(|d|-|y|)\mathrm{sgn}(y)\mathbf{x}^* \\ &= -2[|d|\mathrm{sgn}(y) - y]\mathbf{x}^*, \end{aligned} \quad (7)$$

where $\mathrm{sgn}(y) = y/|y|$. The gradient of $J_{\mathrm{p}}(d,y)$ with respect to $\mathbf{w}^*$ is

$$\begin{aligned}\nabla_{\mathbf{w}^*} J_{\mathrm{p}}(d,y) &= -2|d|[\cos(\angle d - \angle y) \\ &\quad + j\sin(\angle d - \angle y) - 1]\mathrm{sgn}(y)\mathbf{x}^* \\ &= -2[d - |d|\mathrm{sgn}(y)]\mathbf{x}^*. \end{aligned} \quad (8)$$

Our proposed algorithm is defined as

$$\begin{aligned}\mathbf{w}_{k+1} &= \mathbf{w}_k - \frac{\mu_{\mathrm{m},k}}{2}\nabla_{\mathbf{w}^*} J_{\mathrm{m}}(d_k,y_k) \\ &\quad - \frac{\mu_{\mathrm{p},k}}{2}\nabla_{\mathbf{w}^*} J_{\mathrm{p}}(d_k,y_k), \end{aligned} \quad (9)$$

where we have reinserted the time indices of all quantities. Similar to the LMP algorithm, the proposed LMMP algorithm has two (possibly time-varying) step sizes $\mu_{\mathrm{m},k}$ and $\mu_{\mathrm{p},k}$ that control the adaptation behavior. Substituting for the gradient expressions given above, we obtain

$$\begin{aligned}\mathbf{w}_{k+1} &= \mathbf{w}_k + \mu_{\mathrm{m},k}[|d_k|\mathrm{sgn}(y_k) - y_k]\mathbf{x}_k^* \\ &\quad + \mu_{\mathrm{p},k}[d_k - |d_k|\mathrm{sgn}(y_k)]\mathbf{x}_k^* \end{aligned} \quad (10)$$

Finally, we group like terms to get

$$\begin{aligned}\mathbf{w}_{k+1} &= \mathbf{w}_k + [\mu_{\mathrm{p},k} d_k - \mu_{\mathrm{m},k} y_k \\ &\quad + (\mu_{\mathrm{m},k} - \mu_{\mathrm{p},k})|d_k|\mathrm{sgn}(y_k)]\,\mathbf{x}_k^* \quad (11) \\ y_k &= \mathbf{w}_k^T \mathbf{x}_k. \quad (12)\end{aligned}$$

The above equations define the LMMP algorithm. The normalized LMMP algorithm sets $\mu_{\mathrm{m},k} = \mu_{\mathrm{m}}/(\mathbf{x}_k^H\mathbf{x}_k)$ and $\mu_{\mathrm{p},k} = \mu_{\mathrm{p}}/(\mathbf{x}_k^H\mathbf{x}_k)$.

Remark 3: The LMMP algorithm has two different step sizes $\mu_{\mathrm{m},k}$ and $\mu_{\mathrm{p},k}$ at iteration k. The first step size largely controls the adaptation dynamics of the magnitude of the output signal, and the second step size largely controls the adaptation dynamics of the phase of the output signal. The first point is

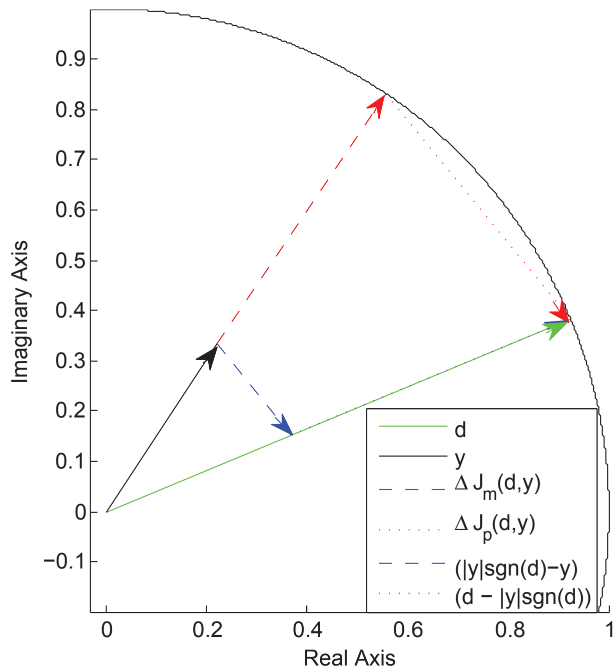

Fig. 1. Various performance improvements for magnitude-phase algorithms as described in Remark 5.

clear given the form of (5). To see the latter point, we can rewrite the form of $\nabla_{\mathbf{w}^*} J_{\mathrm{p}}(d,y)$ in (8) as

$$\begin{aligned}\nabla_{\mathbf{w}^*} J_{\mathrm{p}}(d,y) &= -[d - |d|\exp(j\angle y)]\mathbf{x}^* \quad (13) \\ &= |d|[\exp(j\angle d) - \exp(j\angle y)]\mathbf{x}^*. \quad (14)\end{aligned}$$

The gradient clearly depends on the angle of the output signal and not on its magnitude.

Remark 4: The LMMP algorithm includes several existing algorithms as special cases. Choosing $\mu_{\mathrm{m},k} = \mu_{\mathrm{p},k} = \mu$ yields the CLMS algorithm [1]. Choosing $\mu_{\mathrm{m},k} = \alpha/(\mathbf{x}_k^H\mathbf{x}_k)$ and $\mu_{\mathrm{p},k} = 0$, the *a posteriori* form of the CMCE algorithm in [2] is obtained. Finally, if one considers the alternate form of the algorithm given by

$$\begin{aligned}\mathbf{w}_{k+1} &= \mathbf{w}_k + \mu_{\mathrm{m},k}(d_k - y_k)\mathbf{x}_k^* \\ &\quad + (\mu_{\mathrm{p},k} - \mu_{\mathrm{m},k})\left(d_k - \frac{|d_k|}{|y_k|}y_k\right)\mathbf{x}_k^*, \end{aligned} \quad (15)$$

it can be shown via a Taylor series expansion that

$$\left(d_k - \frac{|d_k|}{|y_k|}y_k\right) \approx (\angle d_k - \angle y_k)\frac{j|d_k||y_k|}{y_k^*}, \quad (16)$$

if $|\angle d_k - \angle y_k|$ is small, yielding the LMP algorithm [3] for the specific step size choices of $\mu_1 = \mu_{\mathrm{m},k}$ and $\mu_2 = (\mu_{\mathrm{p},k} - \mu_{\mathrm{m},k})|d_k||y_k|$.

Remark 5: Like the LMP algorithm of [3], it is useful to look at the geometrical properties of the LMMP algorithm in the complex plane. Fig. 1 shows a vector-based example similar

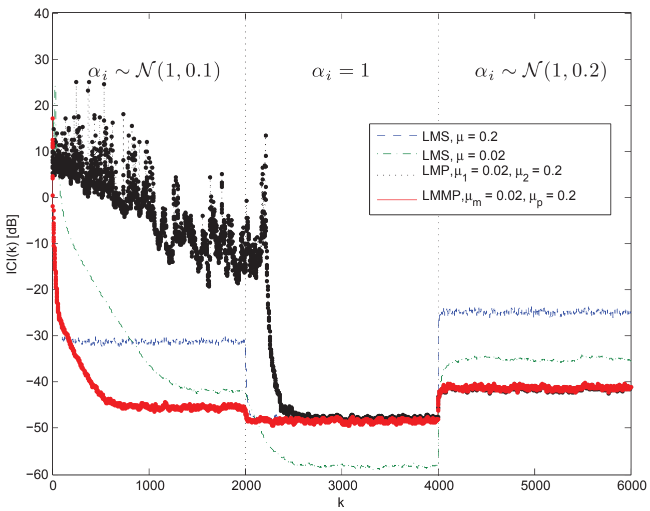

Fig. 2. Evolution of the average ICIs in the antenna array example.

to a figure in [3] illustrating all of the important vectors that describe the LMMP update. Using a similar "performance improvement metric" as introduced in [3],

$$\Delta y \quad = \quad \mathbf{w}_{k+1}^T \mathbf{x} - \mathbf{w}_k^T \mathbf{x}, \tag{17}$$

we obtain for the LMMP algorithm the expression

$$\Delta y_{\text{LMMP}} \;=\; ||\mathbf{x}||^2 \left[\mu_{\text{m}} \Delta J_{\text{m}}(d,y) + \mu_{\text{p}} \Delta J_{\text{p}}(d,y)\right] \tag{18}$$

$$\Delta J_{\text{m}}(d,y) \;=\; |d|\text{sgn}(y) - y \tag{19}$$

$$\Delta J_{\text{p}}(d,y) \;=\; d - |d|\text{sgn}(y) \tag{20}$$

Fig. 1 shows the accumulation of these vectors in the case of normalized step sizes, such that the NLMS algorithm is obtained. The decomposition of the *a posteriori* output is in polar coordinates, in which $\Delta J_{\text{m}}(d,y)$ is in the radial direction of y and $\Delta J_{\text{p}}(d,y)$ performs a rotation on $y + \Delta J_{\text{m}}(d,y)$ to reach d. It is difficult to compare the least mean phase and LMMP algorithms geometrically due to the different step sizes required by the algorithms in practice.

There are an infinite number of ways to decompose $d - y$; one example is shown in Fig. 1 and yields the update

$$\mathbf{w}_{k+1} \;=\; \mathbf{w}_k + \mu_{1,k}[|y_k|\text{sgn}(d_k) - y_k]\mathbf{x}_k^* + \mu_{2,k}[d_k - |y_k|\text{sgn}(d_k)]\mathbf{x}_k^*. \tag{21}$$

While this alternative approach might seem viable, simulations of (21) indicate that it can behave in an unstable manner. The LMP algorithm [3] also can suffer from erratic behavior because of the y_k^* term in the denominator of the phase cost term when μ_2 is constant. From simulations, we observe that the LMMP algorithm is much better behaved from both a convergence and stability standpoint as compared to these other approaches, especially when step size normalization is used.

3. NUMERICAL EXAMPLES

3.1 Array Processing Example: Let $\mathbf{x}_k$ denote individual snapshots from an $L = 6$-element uniform linear array with

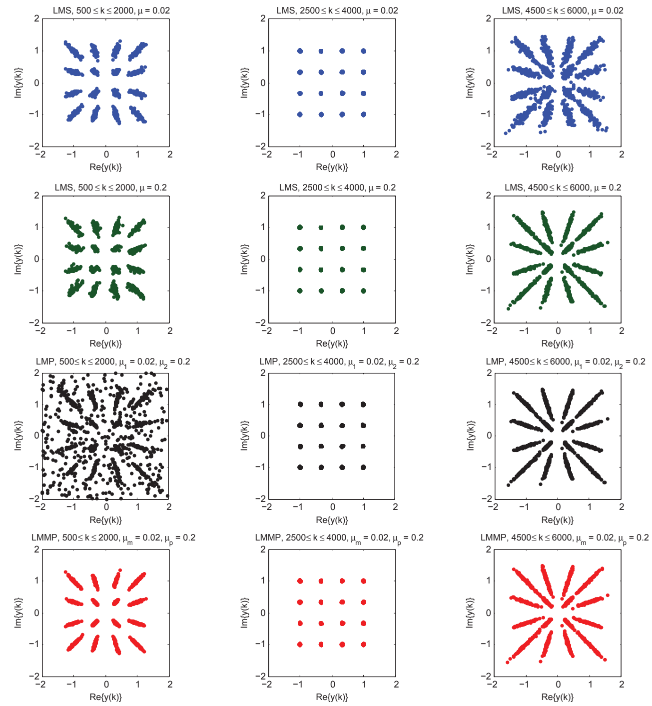

Fig. 3. Signal constellations for the various algorithms in the antenna array example.

$\lambda/4$ spacing. Three signal sources of equal powers are received at arrival angles of $[\phi_1\ \phi_2\ \phi_3] = [-31^\circ\ 11^\circ\ 35^\circ]$ to the array and are corrupted by complex circular Gaussian noise with element powers of 0.0001. Each source is generated from a 16-QAM signal constellation according to the model $s_{i,k} = \alpha_{i,k} d_{i,k}$, where $d_{i,k}$ is the ith i.i.d. 16-QAM symbol sequence satisfying $\max_k \text{Re}\{d_{i,k}\} = \max_k \text{Im}\{d_{i,k}\} = 1$, and $\alpha_{i,k}$ is the unknown time-varying amplitude sequence of the ith transmitter. Such time-varying amplitudes could occur due to signal fading local to the transmitter, imperfections such as nonlinear power amplification, or other effects. The distribution of $\alpha_{i,k}$ is real-valued Gaussian with a mean of one and a variance that changes sequentially every 2000 snapshots from $\sigma_\alpha^2 = 0.1$ to $\sigma_\alpha^2 = 0$ to $\sigma_\alpha^2 = 0.2$. The normalized versions of the LMS, LMP, and LMMP algorithms are used to estimate the first signal source from the snapshots using $d_k = d_{1,k}$. The average interchannel interference (ICI) is found at each snapshot for each algorithm.

Fig. 2 shows the evolution of the average ICI over 1000 different simulation runs for four normalized algorithms: 1) LMS with $\mu = 0.2$, 2) LMS with $\mu = 0.02$, 3) LMP with $\mu_1 = 0.02$ and $\mu_2 = 0.2$, and 4) LMMP with $\mu_{\text{m}} = 0.02$ and $\mu_{\text{p}} = 0.2$. As can be seen, the LMMP algorithm outperforms all other algorithms for the first 2000 snapshots, converging as fast as NLMS with $\mu = 0.2$ but achieving a lower steady-state ICI as compared to either LMS algorithm during this period. When the amplitude fluctuations stop in the second 2000 snapshots, the LMS algorithm with $\mu = 0.02$ converges to a lower steady-state ICI, and the LMMP algorithm

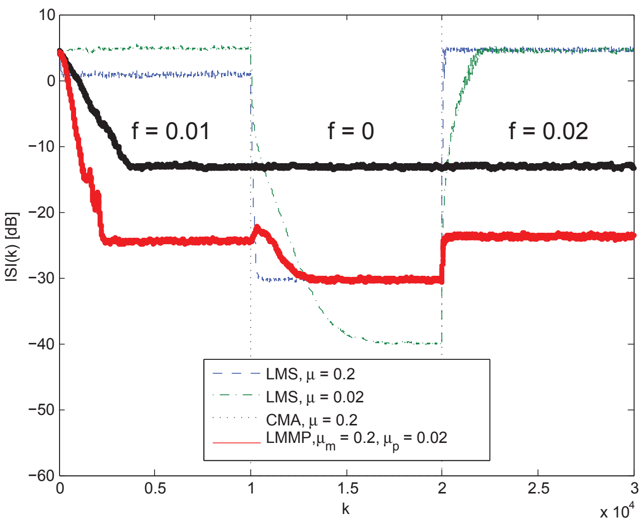

Fig. 4. Evolution of the average ISIs in the channel equalization example.

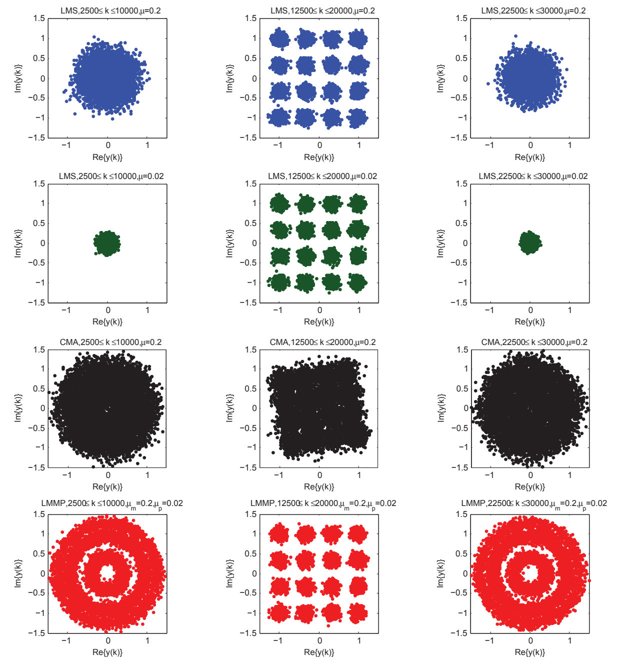

Fig. 5. Signal constellations for the various algorithms in the channel equalization example.

has about the same ICI as that of the LMS algorithm with $\mu = 0.2$. When the amplitude fluctuations commence again in the last 2000 snapshots, the performance of the LMMP algorithm is again superior. The average performance of the LMP and LMMP algorithms are similar when the error is small, but LMMP has better initial convergence than does LMP.

Fig. 3 shows the complex baseband signal constellations produced by each algorithm in this example over the last 1500 snapshots of each adaptation period. The amplitude spread of the output signals cannot be estimated using the chosen desired response signal during the first and third adaptation periods. The LMMP algorithm is the least affected by these amplitude fluctuations.

3.2 Channel Equalization Example: Define the input symbol sequence s_k as an i.i.d. 16-QAM source as before. The received signal x_k is generated as

$$x_k = \eta_k + e^{j2\pi k f_k}[h_0 s_k + h_1 s_{k-1} + h_2 s_{k-2}], \quad (22)$$

where $\{h_0 \; h_1 \; h_2\} = \{0.2e^{j0.1\pi} \; 1e^{j0.2\pi} \; 0.1e^{j0.3\pi}\}$ and the channel noise η_k is complex circular Guassian with variance 0.01. The parameter f_k models frequency offset effects due to physical motion, and its value varies from $f_k = 0.01$ to $f_k = 0$ to $f_k = 0.02$ over three different 10000-sample periods. Four different $L = 11$-tap FIR adaptive equalizers are applied to x_k, with $d_k = s_{k-5}$: 1) LMS with $\mu = 0.2$, 2) LMS with $\mu = 0.02$, 3) the constant modulus algorithm (CMA) with $\mu = 0.2$, and 4) LMMP with $\mu_{\mathrm{m}} = 0.2$ and $\mu_{\mathrm{p}} = 0.02$. The LMP algorithm cannot be tuned to de-emphasize frequency offsets and was not tested. The average intersymbol interference (ISI) is found for each algorithm at each iteration.

Fig. 4 shows the evolution of the average ISI over 100 different simulation runs for the four algorithms. As can be seen, the LMMP algorithm outperforms the other three algorithms over the first 10000 samples, providing the lowest averaged ISI value over this adaptation period. When there is no Doppler, the LMS algorithm with $\mu = 0.02$ eventually produces the lowest ISI, and the performance of the LMMP algorithm approaches that of LMS with $\mu = 0.2$. When Doppler effects are reintroduced for the final 10000 samples, LMMP again provides the lowest ISI.

Fig. 5 shows the complex baseband signal constellations produced by each algorithm in this example over the last 7500 samples of each adaptation period. The Doppler shift cannot be adequately tracked using any of the algorithms. The NLMMP algorithm provides the best performance in such cases by producing a rotated signal constellation during the first and third adaptation periods and accurately estimating the signal constellations during the second adaptation period.

4. CONCLUSIONS

In this paper, the least-mean-magnitude-phase algorithm is introduced for complex-valued adaptive signal processing. Simulations indicate its utility for various communication tasks. A stability analysis of the algorithm is underway.

5. REFERENCES

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

[2] M. Rupp, "On the separation of channel and frequency estimation," *Conf. Record 36th IEEE Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, vol. 2, pp. 1186-1190, Nov. 1998.

[3] A. Tarighat and A.H. Sayed, "Least mean-phase adaptive filters with applications to communications systems," *IEEE Signal Processing Lett.*, vol. 11, no. 2, pp. 220-223, Feb. 2004.