STABILITY ANALYSIS OF THE LEAST-MEAN-MAGNITUDE-PHASE ALGORITHM

Scott C. Douglas[1,2]

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

Danilo P. Mandic

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

ABSTRACT

The least-mean-magnitude-phase (LMMP) algorithm is useful for complex-valued signal processing applications where control of magnitude and/or phase error information is needed to achieve good overall performance. Due to the highly-nonlinear nature of the update terms in the LMMP and related methods, few convergence and stability results exist to guide step size choices for such algorithms. In this paper, we provide a rigorous stability and convergence analysis of the LMMP algorithm using robustness procedures and give sufficient stability conditions on the magnitude and phase step sizes to guarantee contraction-mapping behavior. We also provide an approximate relation for the steady-state MSE as a function of the step size values. Simulations verify the predictive powers of our analytical results and yield useful insights on step size choices in practice.

Index Terms— Adaptive algorithms, adaptive filters, adaptive signal processing, adaptive systems, algorithm design and analysis.

1. INTRODUCTION

The complex least-mean-square (CLMS) algorithm is a simple procedure for adjusting the coefficients of an adaptive system to model linear, possibly time-varying relationships between complex-valued signals. Originally derived in [1], it has found use in numerous applications, particularly in communications system tasks where a baseband signal representation admits a complex-valued system model.

In some situations, however, the CLMS algorithm may not perform adequately, particularly if the signals being used have time-varying behavior that depends on their complex amplitudes or complex phases, such as those caused by Doppler effects [2] or by the inherent phase characteristics of the signals being processed [3]. To this end, various algorithms have been proposed to address the adaptation of a complex linear system model based on amplitude and/or phase information to improve overall performance, including (1) the constant modulus channel estimator (CMCE) algorithm [2], (2) the least-mean-phase (LMP) algorithm [3], and (3) the least-mean-magnitude-phase (LMMP) algorithm [4]. Of these three algorithms, the LMMP algorithm includes the CMCE algorithm as a special case, and it has tunable parameters that allow it to track fast variations in either the amplitude statistics or the phase statistics of the signals being processed. The LMMP algorithm was also shown to outperform the LMP algorithm in cases where signal amplitude variations are significant [4]. The main drawback of these extensions to the CLMS algorithm has been a lack of solid convergence results that guide users as to the choice of step sizes to obtain consistent stable behavior. While the CMCE algorithm has been proven to cause a measure of the magnitude error of the output to converge [2], no relationship to the phase error has been considered. To our knowledge, no rigorous convergence proofs or stability analyses of the LMP or the LMMP algorithms currently exist in the scientific literature.

In this paper, we provide a rigorous stability analysis of the normalized least-mean-magnitude-phase (normalized LMMP) algorithm for complex-valued signals. Our analysis of the normalized LMMP algorithm uses robustness methods [5]–[9] originally developed for the analysis of the LMS algorithm and extended to the complex signal case, and provides useful sufficient conditions on the magnitude step size $\mu_{\rm m}$ and phase step size $\mu_{\rm p}$ to guarantee convergence. We show that the *signal independent* choices

$$0 \; < \; \mu_{\rm m} \; < \; 2 \tag{1}$$

$$\frac{\mu_{\rm m}}{2} \; < \; \mu_{\rm p} \; < \; 1 + \frac{\mu_{\rm m}}{2} \tag{2}$$

are sufficient to cause the normalized LMMP algorithm to impose a contraction on the parameter error for system modeling for ideal (*i.e.* noise-free) desired signals. In the more-realistic case of noisy signal observations, we obtain the following approximate bound between the error sequence e_k and the observation noise sequence η_k for step size choices satisfying (1)–(2), where $\mathbf{x}_k$ is the input signal vector sequence:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{|e_k|^2}{||\mathbf{x}_k||^2} \leq f(\mu_{\rm p}, \mu_{\rm m}) \left[\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{|\eta_k|^2}{||\mathbf{x}_k||^2} \right] \tag{3}$$

$$f(\mu_{\rm p}, \mu_{\rm m}) = \frac{2\left(\mu_{\rm p} + |\mu_{\rm m} - \mu_{\rm p}|\right)}{\mu_{\rm p}(2-\mu_{\rm p}) - 2|\mu_{\rm m} - \mu_{\rm p}|(1-\mu_{\rm p})}. \tag{4}$$

Thus, the complex error produced by the LMMP algorithm

is well-behaved for the step size ranges in (1)–(2), thereby establishing convergence and stability results for this algorithm. Simulations show that these results are reasonable predictors of stability and steady-state behaviors of the LMMP algorithm and provide useful guidance on step size choices.

2. NORMALIZED LEAST-MEAN-MAGNITUDE-PHASE ALGORITHM

The normalized LMMP algorithm adjusts the coefficient vector $\mathbf{w}_k = [w_{1,k} \cdots w_{L,k}]^T$ of a linear adaptive system with L complex coefficients to model a desired response signal d_k via the relation $y_k = \mathbf{w}_k^T \mathbf{x}_k$, where $\mathbf{x}_k = [x_{1,k} \cdots x_{L,k}]^T$ is the complex-valued input signal vector at time k. It attempts to minimize the instantaneous least-mean-square cost $J_{\rm LMS}(d_k, y_k) = |d_k - y_k|^2$ via the decomposition

$$\begin{aligned} J_{\rm LMS}(d_k, y_k) &= J_{\rm m}(d_k, y_k) + J_{\rm p}(d_k, y_k) && (5)\\ J_{\rm m}(d, y) &= (|d| - |y|)^2 && (6)\\ J_{\rm p}(d, y) &= 2|d||y|(1 - \cos(\angle d - \angle y)), && (7) \end{aligned}$$

where (6) and (7) denote the magnitude and phase costs, respectively, and $y = |y|e^{j\angle y}$.

The normalized LMMP algorithm is derived from

$$\begin{aligned} \mathbf{w}_{k+1} &= \mathbf{w}_k - \frac{\mu_{\rm m}}{2||\mathbf{x}_k||^2} \nabla_{\mathbf{w}^*} J_{\rm m}(d_k, y_k) \\ &\quad - \frac{\mu_{\rm p}}{2||\mathbf{x}_k||^2} \nabla_{\mathbf{w}^*} J_{\rm p}(d_k, y_k), \qquad (8) \end{aligned}$$

where $\nabla_{\mathbf{w}^*} J_\cdot(d, y)$ is the complex gradient of $J_\cdot(d, y)$ with respect to $\mathbf{w}^*$ [10] for each respective cost and $||\mathbf{x}||^2 = \mathbf{x}^{*T}\mathbf{x}$. The update relation obtained from this expression is

$$\begin{aligned} \mathbf{w}_{k+1} &= \mathbf{w}_k + \frac{\mu_{\rm m}}{||\mathbf{x}_k||^2}[|d_k|{\rm sgn}(y_k) - y_k]\mathbf{x}_k^* \\ &\quad + \frac{\mu_{\rm p}}{||\mathbf{x}_k||^2}[d_k - |d_k|{\rm sgn}(y_k)]\mathbf{x}_k^*. \qquad (9) \end{aligned}$$

In this algorithm, the magnitude step size $\mu_{\rm m}$ and phase step size $\mu_{\rm p}$ are chosen to control the degrees to which the magnitude cost $J_{\rm m}(d, y)$ and the phase cost $J_{\rm p}(d, y)$ influence the dynamics of the adaptive system. Note that for $\mu_{\rm m} = \mu_{\rm p} = \mu$, the normalized CLMS algorithm is obtained, and for $\mu_{\rm p} = 0$, the *a posteriori* form of the CMCE algorithm is obtained.

3. STABILITY ANALYSIS

The proof of convergence of the normalized LMMP algorithm uses the robustness analysis first used in the analysis of the LMS algorithm and its variants [5, 6] as extended to the complex case. Through this analysis, we prove the sufficient conditions given in (1)–(2) and develop the error bound in (3). This analysis assumes that the desired response is modeled as

$$d_k = \mathbf{w}_{opt}^T \mathbf{x}_k + \eta_k \qquad (10)$$

where $\mathbf{w}_{opt}$ is the optimum coefficient vector and η_k is an observation noise sequence that is statistically-independent of $\mathbf{x}_k$. For this analysis, we write the normalized LMMP algorithm in the following convenient form:

$$\begin{aligned} \mathbf{w}_{k+1} &= \mathbf{w}_k + \frac{2\mu_{\rm p} - \mu_{\rm m}}{||\mathbf{x}_k||^2} c_k e_k \mathbf{x}_k^* && (11)\\ c_k &= \frac{\mu_{\rm p}}{2\mu_{\rm p} - \mu_{\rm m}} - \left(\frac{\mu_{\rm m} - \mu_{\rm p}}{2\mu_{\rm p} - \mu_{\rm m}}\right)\left(\frac{|d_k| - |y_k|}{e_k}\right)\frac{y_k}{|y_k|} && (12) \end{aligned}$$

The reason for this choice can be found from the following bound, which is a result of the reverse triangle inequality:

$$||d_k| - |y_k|| \le |e_k|, \qquad (13)$$

with equality if and only if $d_k = y_k$, *i.e.* $e_k = 0$. Geometrically, this fact can also be verified because the magnitude of $e_k = d_k - y_k$ always increases with increasing phase error for a given $|d_k|$ and $|y_k|$.

To proceed, define

$$\widetilde{\mu} = 2\mu_{\rm p} - \mu_{\rm m}. \qquad (14)$$

We consider the parameter error vector $\mathbf{v}_k = \mathbf{w}_k - \mathbf{w}_{opt}$ and write the normalized LMMP update as

$$\begin{aligned} \mathbf{v}_{k+1} &= \mathbf{v}_k + \frac{\widetilde{\mu}}{||\mathbf{x}_k||^2} c_k e_k \mathbf{x}_k^* && (15)\\ e_k &= \eta_k - \mathbf{v}_k^T \mathbf{x}_k. && (16) \end{aligned}$$

Pre-multiplying both sides of the above relation by their Hermitian transposes, we obtain

$$\begin{aligned} ||\mathbf{v}_{k+1}||^2 &= ||\mathbf{v}_k||^2 - \frac{\widetilde{\mu}}{||\mathbf{x}_k||^2}(c_k + c_k^*)(|e_k|^2 - |\eta_k|^2) \\ &\quad + \frac{\widetilde{\mu}^2}{||\mathbf{x}_k||^2}|c_k|^2 |e_k|^2 \\ &\quad - \frac{\widetilde{\mu}}{||\mathbf{x}_k||^2} 2{\rm Re}\{c_k \eta_k^* \mathbf{v}_k^T \mathbf{x}_k\}, \qquad (17) \end{aligned}$$

where we have used $e_k \eta_k^* = |\eta_k|^2 - \mathbf{v}_k^T \mathbf{x}_k \eta_k^*$ to simplify the relation. At this point, we consider two cases:

Case 1: No observation noise $\eta_k = 0$. In this case, we have

$$e_k = -\mathbf{v}_k^T \mathbf{x}_k, \qquad (18)$$

such that

$$\begin{aligned} &||\mathbf{v}_{k+1}||^2 - ||\mathbf{v}_k||^2 \\ &= \left[-\widetilde{\mu}(c_k + c_k^*) + \widetilde{\mu}^2 |c_k|^2\right] \frac{|e_k|^2}{||\mathbf{x}_k||^2}. \qquad (19) \end{aligned}$$

For a non-zero error signal, this relation is clearly a contraction mapping for $\mathbf{v}_k$ so long as

$$\widetilde{\mu}(c_k + c_k^*) - \widetilde{\mu}^2 |c_k|^2 > 0, \qquad (20)$$

which becomes $|1-\widetilde{\mu}c| \ < \ 1$ or

$$\left|1-\mu_{\rm p}-(\mu_{\rm m}-\mu_{\rm p})\left(\frac{|d_k|-|y_k|}{e_k}\right)\frac{y_k}{|y_k|}\right| \ < \ 1. \qquad (21)$$

Now, for any real or complex a or b, we have

$$|a-b| \leq |a|+|b|, \qquad (22)$$

which when combined with (13) and (21) yields the following three conditions independent of the values of d_k and y_k:

$$|1-\mu_{\rm p}| \ < \ 1 \qquad (23)$$
$$|1-\mu_{\rm p}-(\mu_{\rm m}-\mu_{\rm p})| \ < \ 1 \qquad (24)$$
$$|1-\mu_{\rm p}+(\mu_{\rm m}-\mu_{\rm p})| \ < \ 1. \qquad (25)$$

Simplifying these relations, we obtain the corresponding conditions in (1) and (2).

Case 2: Observation noise $\eta_k \neq 0$. In this case, we apply the iteration in (17) n times and divide both sides by n to obtain

$$\begin{aligned}\frac{1}{n}||\mathbf{v}_n||^2 \ &= \ \frac{1}{n}||\mathbf{v}_0||^2 \\ &\quad -\frac{1}{n}\sum_{k=0}^{n-1}\frac{|e_k|^2}{||\mathbf{x}_k||^2}\left[\mu_{\rm p}\left(c_k+c_k^*\right)-\mu_p^2\,|c_k|^2\right] \\ &\quad +\frac{1}{n}\sum_{k=0}^{n-1}\frac{|\eta_k|^2}{||\mathbf{x}_k||^2}\mu_{\rm p}\left(c_k+c_k^*\right) \\ &\quad -\frac{1}{n}\sum_{k=0}^{n-1}\frac{\mu_{\rm p}}{||\mathbf{x}_k||^2}2{\rm Re}\{c_k\eta_k^*\mathbf{v}_k^T\mathbf{x}_k\}. \qquad (26)\end{aligned}$$

The left-hand side of (26) is always non-negative, and as $n \rightarrow \infty$, the first term on the right-hand side of (26) tends to zero. Moreover, although c_k functionally depends on both $\mathbf{x}_k$ and η_k, in practice the last term of (26) is much smaller than the other terms in this relation, particularly if the input and observation noise signals are symmetric and well-behaved. Thus, we neglect the last term in (26), which yields the inequality

$$\begin{aligned}&\lim_{n\rightarrow\infty}\frac{1}{n}\sum_{k=0}^{n-1}\frac{|e_k|^2}{||\mathbf{x}_k||^2}\left[\mu_{\rm p}\left(c_k+c_k^*\right)-\mu_p^2\,|c_k|^2\right] \\ &\qquad \leq \ \lim_{n\rightarrow\infty}\frac{1}{n}\sum_{k=0}^{n-1}\frac{|\eta_k|^2}{||\mathbf{x}_k||^2}\mu_{\rm p}\left(c_k+c_k^*\right) \qquad (27)\end{aligned}$$

Taking into account the bounds in (1)–(2) and (13) over all possible ranges and choosing the most conservative values in each case, we obtain the result in (3)–(4).

4. NUMERICAL EVALUATIONS

We now provide numerical evidence of the utility of the technical results derived previously. We first consider the sufficient stability conditions in (1)–(2) and show that LMMP exhibits contraction-mapping behavior for a range of step sizes that includes this stability region. We then explore the value of $f(\mu_{\rm p},\mu_{\rm m})$ describing the relative steady-state MSE level observed for the LMMP algorithm for different step size combinations.

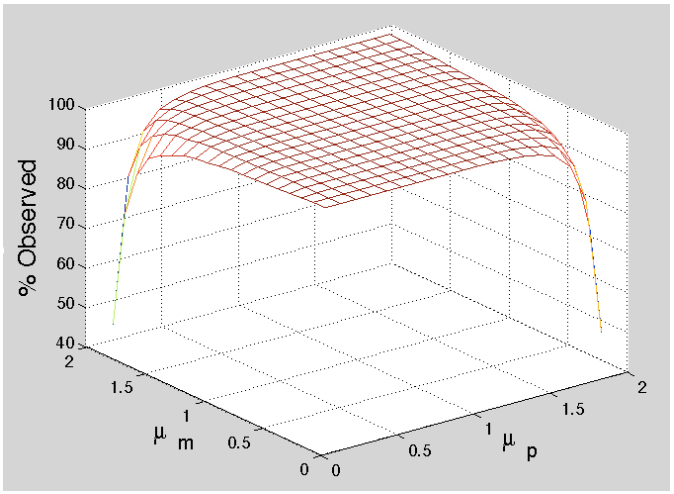

Fig. 1. Percentage of contraction-mapping behaviors observed as a function of step size values $\mu_{\rm p}$ and $\mu_{\rm m}$.

To test the LMMP's stability in an ideal scenario, we apply the algorithm to a nine-coefficient FIR system identification task with zero-mean unit-power uncorrelated complex Gaussian input signals. We perturb $\mathbf{w}_0$ from $\mathbf{w}_{opt}$ by vectors generated from a complex jointly-Gaussian distribution and normalized such that $||\mathbf{v}_0||^2 = 1$. We then apply the LMMP algorithm for a range of step sizes spanning $0.1 \leq \{\mu_{\rm p},\mu_{\rm m}\} \leq 1.9$ and observe the number of times that the convergence of $||\mathbf{v}_k||^2$ is monotonically-decreasing during its entire initial convergence period, determined by either the smaller of $k = 1000$ iterations or the value of k such that $||\mathbf{v}_k||^2 = 10^{-7}$, as calculated across 10000 different Monte Carlo runs for each $\{\mu_{\rm p},\mu_{\rm m}\}$ pair. By representing this fraction of trials as a percentage, we can observe in practice for what step size choices the LMMP algorithm is providing convergent behavior and compare these observations to our sufficient stability conditions for $\mu_{\rm p}$ and $\mu_{\rm m}$.

Fig. 1 shows this observed percentage as a mesh plot across $\mu_{\rm p}$ and $\mu_{\rm m}$. From this figure, we see that the LMMP algorithm provides consistent and stable behavior for a wide range of step sizes, as contraction-mapping behavior is observed for nearly the entire ranges of both $\mu_{\rm p}$ and $\mu_{\rm m}$ over the intervals $(0,2)$. The only step size combinations that yield problematic behavior are when one step size value is chosen close to 2 and the other step size value is near zero, and these combinations are easily avoided.

Fig. 2 shows the observed percentage of contraction-mapping behavior as a quantized grayscale image, where white regions correspond to 100% observations and other regions represent strictly less than 100% of the observations. Also shown is the stability region for $\{\mu_{\rm p},\mu_{\rm m}\}$

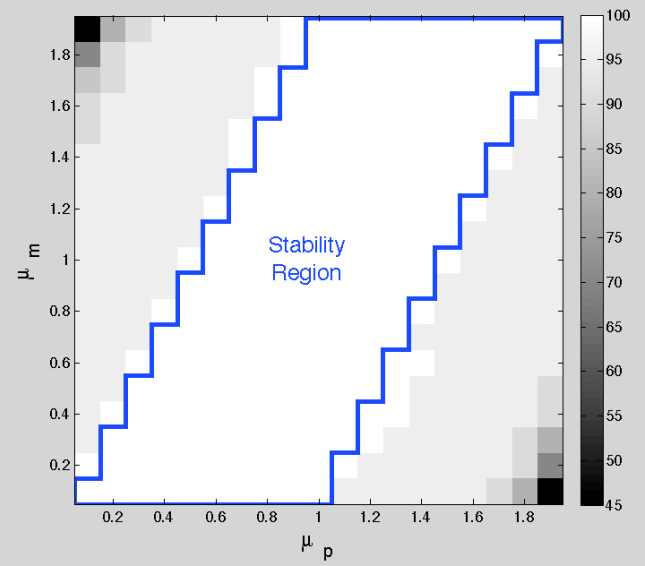

Fig. 2. Percentage of contraction-mapping behaviors observed as compared to the sufficient stability conditions.

from our analysis in blue, adjusted for gridding in this case. As can be seen, our sufficient stability conditions always produce contraction-mapping behavior in our simulations, verifying our results, and the predicted stability region covers the observed contraction-mapping region. The important result from this analysis is the recognition that the LMMP algorithm's stability appears to be well-behaved over $0 < \{\mu_{\rm p}, \mu_{\rm m}\} \leq 1$, which represent the practical values often chosen for the algorithm [4].

Fig. 3 shows the value of $f(\mu_{\rm p}, \mu_{\rm m})$ in dB as a function of $\{\mu_{\rm p}, \mu_{\rm m}\}$ over the stability region in (1)–(2) as a mesh plot (left side) and a contour plot (right side). This figure represents the bound on the relative steady-state MSE produced by LMMP as a function of step size choices. As can be seen, $f(\mu_{\rm p}, \mu_{\rm m})$ has a "boat hull" shape that is generally small but increases significantly in situations where $\mu_{\rm p}$ is small *and* $\mu_{\rm m}$ is either near zero or near $2\mu_{\rm p}$, as well as near $\mu_{\rm p} = \mu_{\rm m} = 2$. This increase in steady-state MSE is also observed for these step size combinations in Monte Carlos simulations. Thus, if a low MSE is important to achieve, one should choose a value of $\mu_{\rm m}$ that is not too small or too large relative to the value of $\mu_{\rm p}$, and one should avoid the upper limits of the stability range for each step size.

To evaluate the behavior of the bound in (3)–(4), we ran simulations of a simple complex-valued nine-coefficient system identification task for different μ_p and μ_m choices and computed time averages of $|e_k|^2/||\mathbf{x}_k||^2$ and $|\eta_k|^2/||\mathbf{x}_k||^2$ from these simulations. To approximate the limiting behavior of the expressions, we used only the last 2500 iterations of these simulations to compute means and σ-error bars for different choices of μ_p and μ_m. Shown in Fig. 4 are the means and σ-error bars using 200 experiments each for the normalized MSEs $|e_k|^2/||\mathbf{x}_k||^2$ and the bound in (3)–(4) computed

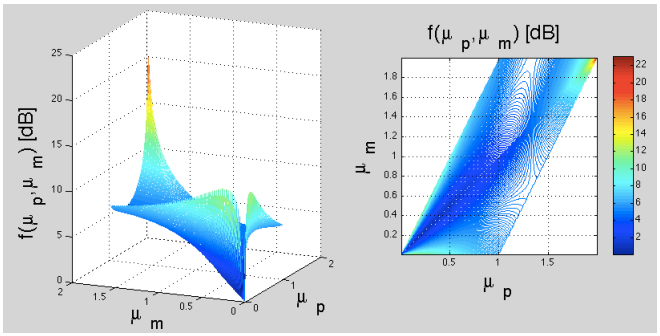

Fig. 3. Value of $f(\mu_{\rm p}, \mu_{\rm m})$ in dB for different step size choices as a mesh plot (left side) and a contour plot (right side).

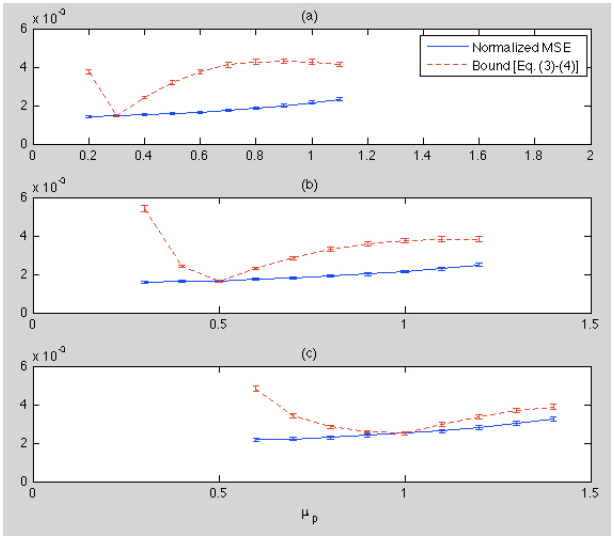

Fig. 4. Numerical evaluation of the bound in (3)-(4) via simulations; see text for explanation.

from $|\eta_k|^2/||\mathbf{x}_k||^2$ as a function of μ_p for (a) $\mu_m = 0.3$, (b) $\mu_m = 0.5$, and (c) $\mu_m = 1.0$, respectively. The bound is clearly tightest when $\mu_p = \mu_m$ corresponding to the complex LMS algorithm, although the upper bound is within a factor of four of the normalized MSE in the cases shown. The bound appears to be most loose for step sizes near $\mu_p = \mu_m/2$, as the LMMP algorithm shows no anomalous behavior in such cases.

5. CONCLUSIONS

This paper provides a simple but useful stability analysis of the least-mean-magnitude-phase (LMMP) algorithm for complex-valued adaptive signal processing applications. The analysis yields practical results on the choice of step sizes for this algorithm, providing step size ranges to guarantee stability. It also yields approximate expressions for the steady-state MSE of the algorithm as a function of the step sizes chosen. Simulations indicate that the results are accurate and provide additional insight on the behavior of this algorithm.

6. 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," *Proc. 32nd 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.

[4] S.C. Douglas and D.P. Mandic, "The least-mean-magnitude-phase algorithm with applications in communications systems," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing,* Prague, Czech Republic, pp. 4152-4155, May 2011.

[5] M. Rupp, "Contraction mapping: An important property in adaptive filters," *Proc. 6th IEEE Digital Signal Processing Workshop,* Yosemite National Park, CA, pp. 273-276, Oct. 1994.

[6] A.H. Sayed and M. Rupp, "Error energy bounds for adaptive gradient algorithms," *IEEE Trans. Signal Processing,* vol. 44, no. 8, pp. 1982-1989, Aug. 1996.

[7] S.C. Douglas and M. Rupp, "A posteriori updates for adaptive filters," *Proc. 31st IEEE Asilomar Conf. Signals, Syst., Comput.,* Pacific Grove, CA, vol. 2, pp. 1641-1645, Nov. 1997.

[8] M. Rupp and S.C. Douglas, "A posteriori analysis of adaptive blind equalizers," *Proc. 32nd Asilomar Conf. Signals, Syst., Comput.,* Pacific Grove, CA, vol. 1, pp. 369-373, Nov. 1998.

[9] A.H. Sayed, *Fundamentals of Adaptive Filtering* (Hoboken, NJ: Wiley, 2003).

[10] D.P. Mandic and S.L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models* (West Sussex, United Kingdom: Wiley, 2009).