Article:
Shi, Long, Zhao, Haiquan and Zakharov, Yuriy orcid.org/0000-0002-2193-4334 (2019) Performance Analysis of Shrinkage Linear Complex-Valued LMS Algorithm. IEEE Signal Processing Letters. pp. 1202-1206. ISSN 1070-9908

https://doi.org/10.1109/LSP.2019.2925957

Performance Analysis of Shrinkage Linear Complex-Valued LMS Algorithm

Long Shi, *Student Member, IEEE,* Haiquan Zhao, *Senior Member, IEEE,* and Yuriy Zakharov, *Senior Member, IEEE*

***Abstract*—The shrinkage linear complex-valued least mean squares (SL-CLMS) algorithm with a variable step-size (VSS) overcomes the conflicting issue between fast convergence and low steady-state misalignment. To the best of our knowledge, the theoretical performance analysis of the SL-CLMS algorithm has not been presented yet. This letter focuses on the theoretical analysis of the excess mean square error (EMSE) transient and steady-state performance of the SL-CLMS algorithm. Simulation results obtained for identification scenarios show a good match with the analytical results.**

***Index Terms*—EMSE, Kronecker product, Rayleigh distribution, shrinkage.**

I. Introduction

THE complex-valued least mean square (CLMS) adaptive filtering algorithm is a well-known estimation technique, which can be considered as an extension of the classical least mean square (LMS) algorithm in the complex domain. It has been successfully applied in the system identification, beamforming and frequency estimation [1]–[5]. As reported in [6], the CLMS algorithm provides good results in the case of circular Gaussian input signals totally described by the covariance matrix, with its pseudo-covariance matrix being zero. In practice, e.g., in communication applications, the complex inputs often have a non-zero pseudo-covariance matrix [7]. To exploit the information of both the matrices, the widely linear CLMS (WL-CLMS) algorithm was proposed [6], [8]. Both the algorithms with time-invariant step-size have been recently analyzed in detail [9]–[12].

For an adaptive filtering algorithm with a fixed step-size, the tradeoff between fast convergence and low steady-state misalignment is unavoidable. To address this issue, the shrinkage linear CLMS (SL-CLMS) algorithm was proposed [13], in which the variable step-size (VSS) is derived by minimizing the energy of the noise-free *a posteriori* error signal.

This letter provides the theoretical analysis of the SL-CLMS algorithm proposed in [13]. By employing properties of the Kronecker product, which is an approach different from the known analysis of complex-valued adaptive algorithms, we arrive at a recursion for computation of the mean-squared error transient and steady-state performance of the algorithm. Simulations for system identification scenarios support the theoretical results.

Notation: The boldface letters denote vectors and matrices. The symbols $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ are, respectively, the complex conjugate, transpose, and Hermitian transpose operators. Symbols $\otimes$, $\max(\cdot)$, and $|\cdot|$ are the Kronecker product, maximum and absolute operators, respectively. The operation $vec(\cdot)$ stacks the matrix into a column. The symbols $E(\cdot)$ and $\mathrm{Tr}(\cdot)$ stand for the mathematical expectation and trace of a matrix, respectively. The symbols $\exp(\cdot)$ and $\mathrm{erf}(\cdot)$ denote the exponential and error functions, respectively. $\boldsymbol{I}_L$ is an $L \times L$ identify matrix.

II. Review of the SL-CLMS Algorithm

Consider a desired signal $d(k)$ at instant k originated from the linear model

$$d(k) = \boldsymbol{w}_o^H \boldsymbol{x}(k) + \eta(k), \tag{1}$$

where $\boldsymbol{w}_o$ denotes the unknown system vector of length L, $\boldsymbol{x}(k) = [x_1(k), x_2(k), \cdots, x_L(k)]^T$ is the input vector, and $\eta(k)$ accounts for the background noise with zero-mean and variance $\sigma_\eta^2 = E[|\eta(k)|^2]$. The error signal $e(k)$ is defined as

$$e(k) = d(k) - \boldsymbol{w}^H(k)\boldsymbol{x}(k), \tag{2}$$

where $\boldsymbol{w}(k)$ is an estimate of $\boldsymbol{w}_o$ at instant k.

In the SL-CLMS algorithm, the weight update is given by

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \mu_k e^*(k)\boldsymbol{x}(k), \tag{3}$$

where μ_k denotes the VSS calculated as [13]

$$\mu_k = \frac{\sigma_{e_a}^2(k)}{E[\|\boldsymbol{x}(k)\|^2]\sigma_e^2(k)}. \tag{4}$$

The quantities $\sigma_e^2(k)$ and $\sigma_{e_a}^2(k)$ are calculated as

$$\sigma_e^2(k) = \lambda\sigma_e^2(k-1) + (1-\lambda)\left|e(k)\right|^2, \tag{5}$$

$$\sigma_{e_a}^2(k) = \lambda\sigma_{e_a}^2(k-1) + (1-\lambda)\left|\hat{e}_a(k)\right|^2, \tag{6}$$

where

$$\hat{e}_a(k) = \mathrm{sign}[e(k)]\max(|e(k)| - t, 0), \tag{7}$$

λ is the forgetting factor ($0 < \lambda \lesssim 1$), $\mathrm{sign}[e(k)] = \frac{e(k)}{|e(k)|}$ and t is a threshold: $t = \sqrt{\theta\sigma_\eta^2/L}$ with $1 \le \theta \le 4$ [13]. In [13],

The work of L. Shi and H. Zhao was partially supported by National Science Foundation of PR China (Grant: 61571374, 61871461, and 61433011), and Doctoral Innovation Fund Program of Southwest Jiaotong University (Grant: D-CX201819). The work of Y. Zakharov was supported by the U.K. EPSRC (Grants EP/P017975/1 and EP/R003297/1).

Long Shi and Haiquan Zhao are with Key Laboratory of Magnetic Suspension Technology and Maglev Vehicle, Ministry of Education, and also with School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, People's Republic of China (e-mail: lshi@my.swjtu.edu.cn; hqzhao_swjtu@126.com). Corresponding author: Haiquan Zhao.

Yuriy Zakharov is with Department of Electronic Engineering, University of York, York YO10 5DD, U.K. (e-mail: yury.zakharov@york.ac.uk).

the quantities $E[\|\boldsymbol{x}(k)\|^2]$ and σ_η^2 are assumed to be known. Note that if the values of $E[\|\boldsymbol{x}(k)\|^2]$ and σ_η^2 are unknown, they can be estimated using estimators proposed in [14], [15].

III. Performance Analysis of the SL-CLMS Algorithm

We make the following assumptions, which are widely used for analyzing VSS adaptive algorithms.

A1: The background noise $\eta(k)$ is zero-mean circular white Gaussian and statistically independent of the noise-free *a priori* error signal $e_a(k) = \tilde{\boldsymbol{w}}^H(k)\boldsymbol{x}(k)$ and input vector $\boldsymbol{x}(k)$, where $\tilde{\boldsymbol{w}} = \boldsymbol{w}(k) - \boldsymbol{w}_o$ is the weight error vector.

A2: The step-size μ_k is statistically independent of the input and weight vectors.

A3: The noise-free *a priori* error signal $e_a(k)$ obeys the zero-mean Gaussian distribution.

Assumption **A1** is one of the most common assumptions in the adaptive filtering theory [1], [16]. Assumption **A2** is widely used for the analysis of VSS adaptive filtering algorithms by considering that the step-size varies slowly, see [17]–[21] and references therein. This assumption might not be very accurate for fast varying step-size, see simulation results below. Assumption **A3** is approximately true when the filter length is large [22], [23].

We define the input covariance matrix $\boldsymbol{R}$ and pseudo-covariance matrix $\boldsymbol{P}$ as

$$\boldsymbol{R} = E[\boldsymbol{x}(k)\boldsymbol{x}^H(k)], \quad \boldsymbol{P} = E[\boldsymbol{x}(k)\boldsymbol{x}^T(k)]. \tag{8}$$

For the weight error vector $\tilde{\boldsymbol{w}}(k)$, from (3) we obtain

$$\tilde{\boldsymbol{w}}(k+1) = \left[\boldsymbol{I}_L - \mu_k \boldsymbol{x}(k)\boldsymbol{x}^H(k)\right]\tilde{\boldsymbol{w}}(k) + \mu_k \eta^*(k)\boldsymbol{x}(k). \tag{9}$$

Post-multiplying (9) by its Hermitian transpose, we arrive at

$$\begin{aligned}
\tilde{\boldsymbol{w}}(k+1)\tilde{\boldsymbol{w}}^H(k+1) &= \tilde{\boldsymbol{w}}(k)\tilde{\boldsymbol{w}}^H(k) \\
&- \mu_k \tilde{\boldsymbol{w}}(k)\tilde{\boldsymbol{w}}^H(k)\boldsymbol{x}(k)\boldsymbol{x}^H(k) \\
&- \mu_k \boldsymbol{x}(k)\boldsymbol{x}^H(k)\tilde{\boldsymbol{w}}(k)\tilde{\boldsymbol{w}}^H(k) \\
&+ \mu_k^2 \boldsymbol{x}(k)\boldsymbol{x}^H(k)\tilde{\boldsymbol{w}}(k)\tilde{\boldsymbol{w}}^H(k)\boldsymbol{x}(k)\boldsymbol{x}^H(k) \\
&+ \mu_k^2 \boldsymbol{x}(k)\boldsymbol{x}^H(k)\left|\eta(k)\right|^2 + \mu_k \tilde{\boldsymbol{w}}(k)\boldsymbol{x}^H(k)\eta(k) \\
&- \mu_k^2 \boldsymbol{x}(k)\boldsymbol{x}^H(k)\tilde{\boldsymbol{w}}(k)\boldsymbol{x}^H(k)\eta(k) + \mu_k \eta^*(k)\boldsymbol{x}(k)\tilde{\boldsymbol{w}}^H(k) \\
&- \mu_k^2 \eta^*(k)\boldsymbol{x}(k)\tilde{\boldsymbol{w}}^H(k)\boldsymbol{x}(k)\boldsymbol{x}^H(k).
\end{aligned} \tag{10}$$

Taking the expectation of (10) and applying assumptions **A1** and **A2** leads to

$$\begin{aligned}
\boldsymbol{Q}(k+1) &= \boldsymbol{Q}(k) - E(\mu_k)[\boldsymbol{R}\boldsymbol{Q}(k) + \boldsymbol{Q}(k)\boldsymbol{R}] + E(\mu_k^2)\sigma_\eta^2\boldsymbol{R} \\
&+ E(\mu_k^2)(\boldsymbol{R}\boldsymbol{Q}(k)\boldsymbol{R} + \boldsymbol{P}\boldsymbol{Q}^*(k)\boldsymbol{P}^* + \boldsymbol{R}\mathrm{Tr}[\boldsymbol{R}\boldsymbol{Q}(k)]),
\end{aligned} \tag{11}$$

where $\boldsymbol{Q}(k) = E[\tilde{\boldsymbol{w}}(k)\tilde{\boldsymbol{w}}^H(k)]$, and the fourth order moment in (10) is decomposed by employing the Gaussian moment factorizing theorem [24]

$$\begin{aligned}
&E[\boldsymbol{x}(k)\boldsymbol{x}^H(k)\tilde{\boldsymbol{w}}(k)\tilde{\boldsymbol{w}}^H(k)\boldsymbol{x}(k)\boldsymbol{x}^H(k)] \\
&\quad = \boldsymbol{R}\boldsymbol{Q}(k)\boldsymbol{R} + \boldsymbol{P}\boldsymbol{Q}^*(k)\boldsymbol{P}^* + \boldsymbol{R}\mathrm{Tr}[\boldsymbol{R}\boldsymbol{Q}(k)].
\end{aligned} \tag{12}$$

Before further proceeding, we make the following approximation [25], [26]:

$$E(\mu_k^2) \approx [E(\mu_k)]^2. \tag{13}$$

This approximation is valid due to the averaging in (5) and (6) for estimates $\sigma_{e_a}^2(k)$ and $\sigma_e^2(k)$. Our numerical analysis (not presented here), for scenarios in Section IV, has shown that this approximation is very accurate. Using (13) in (11), we obtain

$$\begin{aligned}
\boldsymbol{Q}(k+1) &= \boldsymbol{Q}(k) - E(\mu_k)[\boldsymbol{R}\boldsymbol{Q}(k) + \boldsymbol{Q}(k)\boldsymbol{R}] + [E(\mu_k)]^2\sigma_\eta^2\boldsymbol{R} \\
&+ [E(\mu_k)]^2(\boldsymbol{R}\boldsymbol{Q}(k)\boldsymbol{R} + \boldsymbol{P}\boldsymbol{Q}^*(k)\boldsymbol{P}^* + \boldsymbol{R}\mathrm{Tr}[\boldsymbol{R}\boldsymbol{Q}(k)]).
\end{aligned} \tag{14}$$

A. Mean Square Transient Behavior

For arbitrary matrices $\{\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}\}$ of compatible dimensions, $vec(\boldsymbol{X}\boldsymbol{Y}\boldsymbol{Z}) = (\boldsymbol{Z}^T \otimes \boldsymbol{X})vec(\boldsymbol{Y})$ and $\mathrm{Tr}(\boldsymbol{X}\boldsymbol{Y}) = (vec(\boldsymbol{X}^T))^T vec(\boldsymbol{Y})$ [27]. By applying these operations to (14), we arrive at

$$\begin{aligned}
vec(\boldsymbol{Q}(k+1)) &= vec(\boldsymbol{Q}(k)) - E(\mu_k)[(\boldsymbol{I}\otimes\boldsymbol{R})vec(\boldsymbol{Q}(k)) \\
&+ (\boldsymbol{R}^T\otimes\boldsymbol{I})vec(\boldsymbol{Q}(k))] + E(\mu_k^2)\sigma_\eta^2 vec(\boldsymbol{R}) \\
&+ E(\mu_k^2)[(\boldsymbol{R}^T\otimes\boldsymbol{R})vec(\boldsymbol{Q}(k)) + (\boldsymbol{P}^H\otimes\boldsymbol{P})vec(\boldsymbol{Q}^*(k)) \\
&+ vec(\boldsymbol{R})(vec(\boldsymbol{R}^T))^T vec(\boldsymbol{Q}(k))].
\end{aligned} \tag{15}$$

The recursion in (15) can be computed as long as the mean step-size $E(\mu_k)$ is available.

Taking the expectation of (4) and applying **A1**, we obtain

$$E(\mu_k) = \frac{E[\sigma_{e_a}^2(k)]}{E[\|\boldsymbol{x}(k)\|^2]E[\sigma_e^2(k)]}, \tag{16}$$

where

$$E[\sigma_e^2(k)] = \lambda E[\sigma_e^2(k-1)] + (1-\lambda)E[|e(k)|^2], \tag{17}$$

$$E[\sigma_{e_a}^2(k)] = \lambda E[\sigma_{e_a}^2(k-1)] + (1-\lambda)E[|\hat{e}_a(k)|^2]. \tag{18}$$

Here, we have also used the first-order approximation: $E\left\{\frac{\sigma_{e_a}^2(k)}{\sigma_e^2(k)}\right\} \approx \frac{E[\sigma_{e_a}^2(k)]}{E[\sigma_e^2(k)]}$. Note that a more accurate second-order approximation $E\left\{\frac{\sigma_{e_a}^2(k)}{\sigma_e^2(k)}\right\} \approx \gamma\frac{E[\sigma_{e_a}^2(k)]}{E[\sigma_e^2(k)]}$ requires computing the factor $\gamma = 1-\epsilon = 1 - \frac{\mathrm{cov}(\sigma_{e_a}^2(k),\sigma_e^2(k))}{E[\sigma_{e_a}^2(k)]E[\sigma_e^2(k)]} + \frac{\mathrm{var}(\sigma_e^2(k))}{E[\sigma_e^2(k)]^2}$, where $\mathrm{cov}(\cdot)$ denotes the covariance, and $\mathrm{var}(\cdot)$ is the variance [28], [29]. However, our numerical analysis (not presented here), has shown that, for all simulation scenarios in Section IV, $\epsilon << 1$. Therefore, the first-order approximation is used. Note that this approximation is often used for analysis of adaptive filtering algorithms [20], [25], [26].

In (16), the quantity $E[\|\boldsymbol{x}(k)\|^2]$ is available since we have assumed that the input power is known. The recursion for $E[\sigma_e^2(k)]$ is based on $E[|e(k)|^2]$ which is given by

$$E[|e(k)|^2] = \sigma_\eta^2 + \mathrm{Tr}(\boldsymbol{R}\boldsymbol{Q}(k)). \tag{19}$$

The difficulty is the calculation of $E[|\hat{e}_a(k)|^2]$ in (18). By using (7), $E[|\hat{e}_a(k)|^2]$ is expressed as

$$E[|\hat{e}_a(k)|^2] = E\{[\max(|e(k)| - t, 0)]^2\}. \tag{20}$$

Since $e(k) = e_a(k) + \eta(k)$, with assumptions **A1** and **A3**, we obtain that the error $e(k)$ obeys the zero-mean Gaussian distribution. We further assume that the variance of the real and imaginary parts of $e(k)$ have the same variance; this approximation is verified in our simulation in Section IV.

Then, $z = |e(k)|$ obeys the Rayleigh distribution [30] with the probability density function

$$f(z) = \frac{z}{\sigma^2(k)} \exp\left(-\frac{z^2}{2\sigma^2(k)}\right), \quad z \geq 0, \tag{21}$$

where $\sigma^2(k)$ is the variance of the real (imaginary) part of $e(k)$ [30], i.e.,

$$\sigma^2(k) = \frac{E[|e(k)|^2]}{2} = \frac{\sigma_\eta^2 + \mathrm{Tr}(\boldsymbol{R}\boldsymbol{Q}(k))}{2}. \tag{22}$$

From (20) and (21), we have

$$E[|\hat{e}_a(k)|^2] = \frac{1}{\sigma^2(k)} \int_t^\infty (z-t)^2 z \exp\left(-\frac{z^2}{2\sigma^2(k)}\right) \mathrm{d}z. \tag{23}$$

By taking the integral in (23), we arrive at

$$E[|\hat{e}_a(k)|^2] = \Omega_1 - \Omega_2 + \Omega_3, \tag{24}$$

where

$$\begin{aligned}\Omega_1 &= \frac{1}{\sigma^2(k)} \int_t^\infty z^3 \exp\left(-\frac{z^2}{2\sigma^2(k)}\right) \mathrm{d}z \\ &= t^2 \exp\left(-\frac{t^2}{2\sigma^2(k)}\right) + 2\sigma^2(k) \exp\left(-\frac{t^2}{2\sigma^2(k)}\right)\end{aligned} \tag{25}$$

$$\begin{aligned}&\Omega_2 = \frac{1}{\sigma^2(k)} 2t \int_t^\infty z^2 \exp\left(-\frac{z^2}{2\sigma^2(k)}\right) \mathrm{d}z = \\ &2t\left[t \exp\left(-\frac{t^2}{2\sigma^2(k)}\right) - \frac{\sqrt{\pi\sigma^2(k)}}{\sqrt{2}}\left[\mathrm{erf}\left(\frac{t}{\sqrt{2\sigma^2(k)}}\right) - 1\right]\right]\end{aligned} \tag{26}$$

and

$$\begin{aligned}\Omega_3 &= \frac{1}{\sigma^2(k)} t^2 \int_t^\infty z \exp\left(-\frac{z^2}{2\sigma^2(k)}\right) \mathrm{d}z \\ &= t^2 \exp\left(-\frac{t^2}{2\sigma^2(k)}\right).\end{aligned} \tag{27}$$

Based on the above derivation, using (16) – (27), the mean step-size $E(\mu_k)$ is calculated, which is then used in the recursive update (15) to compute the excess mean square error (EMSE) according to

$$\mathrm{EMSE}(k) = (vec(\boldsymbol{R}^T))^T vec(\boldsymbol{Q}(k)). \tag{28}$$

B. Mean Square Steady-state Behavior

As $k \to \infty$ from (15), we obtain the steady-state equation

$$\begin{aligned}&E(\mu_\infty)[(\boldsymbol{I} \otimes \boldsymbol{R})vec(\boldsymbol{Q}(\infty)) + (\boldsymbol{R}^T \otimes \boldsymbol{I})vec(\boldsymbol{Q}(\infty))] \\ &- [E(\mu_\infty)]^2[(\boldsymbol{R}^T \otimes \boldsymbol{R})vec(\boldsymbol{Q}(\infty)) \\ &+ vec(\boldsymbol{R})(vec(\boldsymbol{R}^T))^T vec(\boldsymbol{Q}(\infty))] \\ &= [E(\mu_\infty)]^2 \sigma_\eta^2 vec(\boldsymbol{R}) + [E(\mu_\infty)]^2(\boldsymbol{P}^H \otimes \boldsymbol{P})vec(\boldsymbol{Q}^*(\infty)).\end{aligned} \tag{29}$$

Rearranging (29) results in

$$\begin{aligned}&vec(\boldsymbol{Q}^*(\infty)) = \\ &\quad \boldsymbol{\Psi}_1^{-1} E(\mu_\infty^2)[\sigma_\eta^2 vec(\boldsymbol{R}^*) + (\boldsymbol{P}^T \otimes \boldsymbol{P}^*)vec(\boldsymbol{Q}(\infty))],\end{aligned} \tag{30}$$

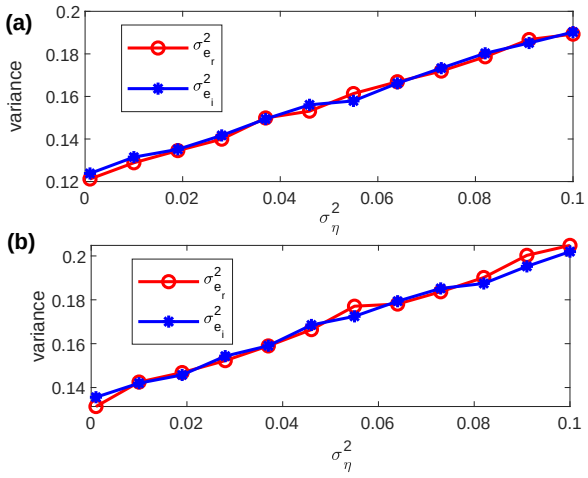

Fig. 1. Evolutions of $\sigma_{e_r}^2$ and $\sigma_{e_i}^2$ for different σ_η^2, $\lambda = 0.95$ and $\theta = 3$. (a) independent Gaussian input; (b) correlated input.

where

$$\begin{aligned}\boldsymbol{\Psi}_1 &= E(\mu_\infty)[\boldsymbol{I} \otimes \boldsymbol{R}^* + \boldsymbol{R}^H \otimes \boldsymbol{I}] \\ &\quad - E(\mu_\infty^2)[\boldsymbol{R}^H \otimes \boldsymbol{R}^* + vec(\boldsymbol{R}^*)(vec(\boldsymbol{R}^T))^H].\end{aligned} \tag{31}$$

Substituting (30) into (29), after some algebra, we arrive at

$$\begin{aligned}&vec(\boldsymbol{Q}(\infty)) = \\ &E(\mu_\infty^2)\boldsymbol{\Psi}_2^{-1}[\sigma_\eta^2 vec(\boldsymbol{R}) + E(\mu_\infty^2)(\boldsymbol{P}^H \otimes \boldsymbol{P})\Psi_1^{-1}\sigma_\eta^2 vec(\boldsymbol{R}^*)],\end{aligned} \tag{32}$$

where

$$\begin{aligned}\boldsymbol{\Psi}_2 &= E(\mu_\infty)[\boldsymbol{I} \otimes \boldsymbol{R} + \boldsymbol{R}^T \otimes \boldsymbol{I}] \\ &\quad - E(\mu_\infty^2)[\boldsymbol{R}^T \otimes \boldsymbol{R} + vec(\boldsymbol{R})(vec(\boldsymbol{R}^T))^T] \\ &\quad - (E(\mu_\infty^2))^2(\boldsymbol{P}^H \otimes \boldsymbol{P})\boldsymbol{\Psi}_1^{-1}(\boldsymbol{P}^T \otimes \boldsymbol{P}^*).\end{aligned} \tag{33}$$

In the steady-state, we can assume that in (19) $\mathrm{Tr}(\boldsymbol{R}\boldsymbol{Q}(k)) << \sigma_\eta^2$, and thus $E[|e(k)|^2] \approx \sigma_\eta^2$ [31]. The steady-state step-size $E(\mu_\infty)$ is calculated using (16)-(18) and (24)-(27). Finally, the steady-state EMSE can be deduced from (28).

IV. Simulation Results

To evaluate our theoretical analysis, we consider system identification scenarios with the 16×1 system vector $\boldsymbol{w}_o = [\boldsymbol{\omega},\ \boldsymbol{\omega},\ \boldsymbol{\omega},\ \boldsymbol{\omega}]^T$, where $\boldsymbol{\omega} = [0.25 + 0.1i,\ 0.5 + 0.75i,\ 0.75 + 0.5i,\ 0.1 + 0.25i]$. The independent Gaussian input is zero-mean non-circular with variance $E[|x(k)|^2] = 1$ and complementary variance $E[x^2(k)] = 0.5$ [11]. The correlated inputs are generated by filtering the independent Gaussian sequence through a first-order auto-regressive model $H(z) = 1/(1 - 0.3z^{-1})$. The background noise is zero-mean circular white Gaussian. The normalized EMSE (NEMSE) $|e_a(k)|^2/\sigma_\eta^2$ is used to evaluate the algorithm performance in Fig. 3 and Fig. 4, while in Fig.5, the EMSE $|e_a(k)|^2$ is shown; all results are obtained by averaging over 1000 simulation trials.

We first present in Fig. 1 variances of real $\sigma_{e_r}^2$ and imaginary $\sigma_{e_i}^2$ parts of $e(k)$. As can be seen, $\sigma_{e_r}^2 \approx \sigma_{e_i}^2$ for all values of the noise variance σ_η^2. This justifies the assumption that

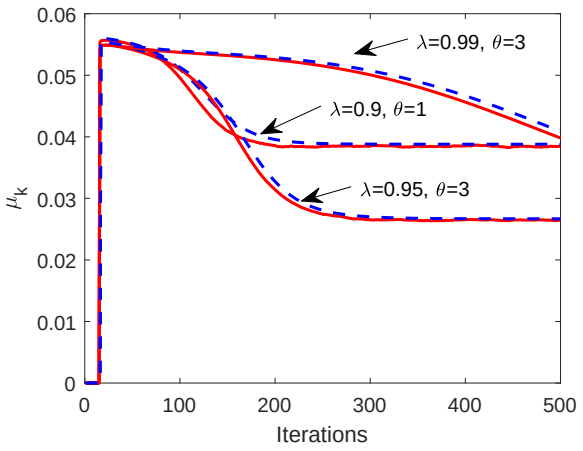

Fig. 2. Evolutions of the step-size for different λ and θ, and $\sigma_\eta^2 = 0.01$. The correlated signal is used as the input. Solid lines: simulation results; dashed lines: theoretical results.

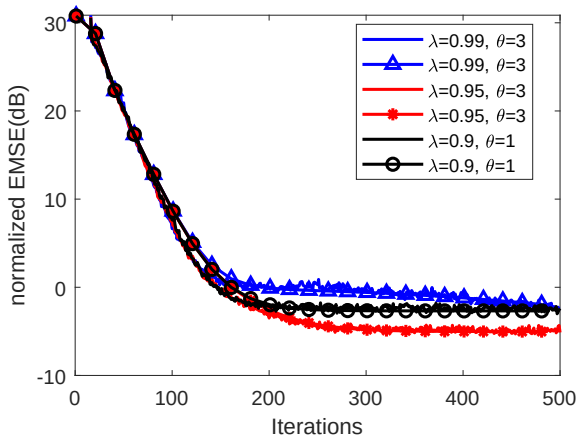

Fig. 4. Normalized EMSE for different values of λ and θ, and $\sigma_\eta^2 = 0.01$. The correlated signal is used as the input. Lines without marks: simulation results; lines with marks: theoretical results.

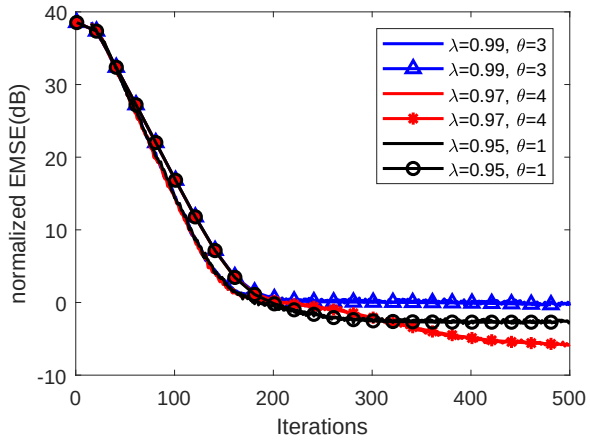

Fig. 3. Normalized EMSE for different values of λ and θ, and $\sigma_\eta^2 = 0.001$. The independent Gaussian signal is used as the input. Lines without marks: simulation results; lines with marks: theoretical results.

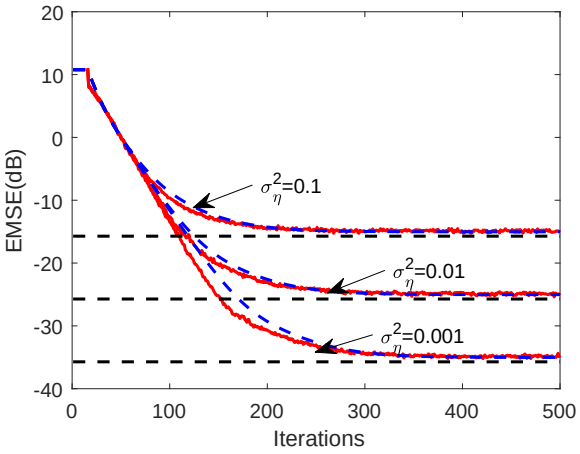

Fig. 5. EMSE for different noise variances; $\lambda = 0.95$ and $\theta = 3$. The correlated signal is used as the input. Red lines: simulation results; blue lines: theoretical transient results; black lines: theoretical steady-state results.

$|e(k)|$ has the Rayleigh distribution, as used in our theoretical analysis.

Fig. 2 shows the evolution of the step-size with iterations for different values of the forgetting factor λ and threshold parameter θ. It is seen that the theoretical prediction is accurate in all the cases, apart from the transient period when the step-size varies very quickly.

Fig. 3 shows the NEMSE for the case of the independent Gaussian input, obtained for different values of λ and θ in the simulation and theoretically predicted. It can be seen that the theoretical prediction is very accurate for all sets of the parameters at all iterations. There is, however, some discrepancy in the transient period due to the fast variation of the step-size.

Fig. 4 presents similar results for the case of the correlated Gaussian input, and again the theoretical prediction is very accurate.

Fig. 5 compares the simulated and theoretical EMSE for different noise variances. For all the noise variances, the theoretical analysis provides good prediction of the steady-state EMSE. When $\sigma_\eta^2 = 0.1$ and $\sigma_\eta^2 = 0.01$, the transient behaviour is also accurately approximated by the theoretical curve. Only for a low noise variance ($\sigma_\eta^2 = 0.001$), there is some deviation between the simulated and theoretical transient EMSE. This deviation is due to the limited accuracy of the approximation in (16).

V. Conclusion

In this letter, we have presented the theoretical analysis of the transient and steady-state EMSE performance of the SL-CLMS adaptive algorithm for the case of non-circular input signal and circular Gaussian noise. Comparison of simulation and theoretical results for identification scenarios with different parameters have shown that the theoretical prediction provided by our analysis is very accurate.

REFERENCES

[1] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York, NY, USA: Wiley, 2003.

[2] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive antenna systems," *J. Acoust. Soc. Amer.*, vol. 42, no. 5, pp. 1175–1176, Dec. 2005.

[3] S. Hossain, M. T. Islam, and S. Serikawa, "Adaptive beamforming algorithms for smart antenna systems," in *Proc. Int. Conf. Control, Autom., Syst. (ICCAS)*, pp. 412–416, Oct. 2008.

[4] Y. Xia, S. C. Douglas, and D. P. Mandic, "Adaptive frequency estimation in smart grid applications: Exploiting noncircularity and widely linear adaptive estimators," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 1812–1816, Jul. 2005.

[5] Y. Xia and D. P. Mandic, "Widely linear adaptive frequency estimation of unbalanced three-phase power systems," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 1, pp. 74–83, Jan. 2012.

[6] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Processing.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[7] B. Picinbono, "On circularity," *IEEE Trans. Signal Processing.*, vol. 42, no. 12, p. 3473–3482, Dec. 1994.

[8] J. Navarro-Moreno, J. Moreno-Kayser, R. Fernandez-Alcala, and J. Ruiz-Molina, "Widely linear estimation algorithms for second-order stationary signals," *IEEE Trans. Signal Processing.*, vol. 57, no. 12, p. 4930–4935, Dec. 2009.

[9] S. C. Douglas and D. P. Mandic, "Performance analysis of the conventional complex LMS and augmented complex LMS algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 3794–3797, Mar. 2010.

[10] D. P. Mandic, S. Kanna, and S. C. Douglas, "Mean square analysis of the CLMS and ACLMS for non-circular signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 3531–3535, Apr. 2015.

[11] Y. Xia and D. P. Mandic, "Complementary mean square analysis of augmented CLMS for second-order noncircular gaussian signals," *IEEE Signal Process. Lett.*, vol. 24, no. 9, pp. 1413–1417, Sep. 2017.

[12] ——, "A full mean square analysis of CLMS for second-order noncircular inputs," *IEEE Trans. Signal Process.*, vol. 65, no. 21, pp. 5578–5590, Nov. 2017.

[13] Y.-M. Shi, L. Huang, C. Qian, and H. C. So, "Shrinkage linear and widely linear complex-valued least mean squares algorithms for adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 119–131, Jan. 2015.

[14] M. A. Iqbal and S. L. Grant, "Novel variable step size NLMS algorithms for echo cancellation," in *IEEE International Conference on Acoustics, Speech and Signal Processing*. Las Vegas, NV, 2008, pp. 241–244.

[15] H.-C. Huang and J. Lee, "A new variable step-size NLMS algorithm and its performance analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 2055–2060, 2012.

[16] L. Shi, H. Zhao, and Y. Zakharov, "Generalized variable step size continuous mixed p-norm adaptive filtering algorithm," *IEEE Trans. Circuits Syst. II, Exp. Briefs.*, Doi: 10.1109/TCSII.2018.2873254, 2018.

[17] A. I. Sulyman and A. Zerguine, "Convergence and steady-state analysis of a variable step-size NLMS algorithm," *Signal Process.*, vol. 83, no. 6.

[18] H. Lee, S. Kim, J. Lee, and W. Song, "A variable step-size diffusion LMS algorithm for distributed estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 7.

[19] R. H. Kwong and E. W. Johnston, "A variable step size LMS algorithm," *IEEE Transactions on signal processing*, vol. 40, no. 7, pp. 1633–1642, 1992.

[20] S. Koike, "A class of adaptive step-size control algorithms for adaptive filters," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1315–1326, 2002.

[21] V. J. Mathews and Z. Xie, "A stochastic gradient adaptive filter with gradient adaptive step size," *IEEE transactions on Signal Processing*, vol. 41, no. 6, pp. 2075–2087, 1993.

[22] S. Zhang, W. Zheng, and J. Zhang, "A new combined-step-size normalized least mean square algorithm for cyclostationary inputs," *Signal Process.*, vol. 141, pp. 261–272, Dec. 2017.

[23] L. Shi, H. Zhao, W. Wang, and L. Lu, "Combined regularization parameter for normalized LMS algorithm and its performance analysis," *Signal Processing*, Doi: https://doi.org/10.1016/j.sigpro.2019.04.014, 2019.

[24] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: analysis and simulations," *IEEE Trans. Signal Process.*, vol. 45, no. 3.

[25] H.-S. Lee, S.-E. Kim, J.-W. Lee, and W.-J. Song, "A variable step-size diffusion LMS algorithm for distributed estimation." *IEEE Trans. Signal Processing*, vol. 63, no. 7, pp. 1808–1820, 2015.

[26] S. Koike, "Analysis of adaptive filters using normalized signed regressor LMS algorithm," *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2710–2723, 1999.

[27] L. Shi and H. Zhao, "Diffusion leaky zero attracting least mean square algorithm and its performance analysis," *IEEE Access.*, vol. 6, pp. 56 911–56 923, 2018.

[28] K. Wolter, *Introduction to variance estimation*. Springer Science & Business Media, 2007.

[29] J. Hayya, D. Armstrong, and N. Gressis, "A note on the ratio of two normally distributed variables," *Management Science*, vol. 21, no. 11, pp. 1338–1341, 1975.

[30] G. L. Stüber, *Principles of mobile communication*. Springer, 1996, vol. 2.

[31] S. Zhang and J. Zhang, "New steady-state analysis results of variable step-size LMS algorithm with different noise distributions," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 653–657, Jun. 2014.