A Full Mean Square Analysis of CLMS for Second-Order Noncircular Inputs

Yili Xia, *Member, IEEE*, and Danilo P. Mandic, *Fellow, IEEE*

***Abstract*—A full mean square transient and steady-state analysis of the complex least mean square (CLMS) algorithm is provided for strictly linear estimation of general second-order noncircular (improper) Gaussian inputs. To this end, we also consider the performance assessment in terms of the evolution of the complementary mean square error (CMSE) and the complementary covariance (pseudocovariance) matrix of the weight error vector of CLMS. This makes it possible to measure the degrees of noncircularity of the output error and the weight error vector, which arise due to second-order noncircularity (improperness) of the system input and system noise. The recently introduced approximate uncorrelating transform, which allows for joint direct diagonalization of both the input covariance and complementary covariance matrices with a single singular value decomposition, is then employed in order to derive a unified bound on the step-size, which guarantees the convergence of both the standard MSE and the proposed CMSE. A joint consideration of the standard mean square performance analysis and the proposed complementary performance analysis is shown to provide full second order, closed form, statistical descriptions of both the transient and steady state performances of CLMS for second-order noncircular (improper) Gaussian input data. Simulations in the system identification setting support the analysis.**

***Index Terms*—Complex LMS (CLMS), second order noncircularity (improperness), complementary mean square convergence analysis, approximate uncorrelating transform (AUT).**

I. Introduction

THE complex least mean square (CLMS) algorithm is probably the most-often-used adaptive signal processing algorithm in the complex domain $\mathbb{C}$, the applications of which include communications, image processing, speech processing, and medicine [1]. For an $N \times 1$ weight vector $\mathbf{h}(k) = [h_1(k), h_2(k), \ldots, h_N(k)]^T$ at a time instant k, the CLMS algorithm updates a simple quadratic error cost function using stochastic gradient descent, and is given by

$$e(k) = d(k) - \mathbf{h}^H(k)\mathbf{x}(k) \tag{1}$$

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \mu e^*(k)\mathbf{x}(k) \tag{2}$$

where $\mathbf{x}(k) = [x_1(k), x_2(k), \ldots, x_N(k)]^T \in \mathbb{C}^{N\times 1}$ is the input vector, $e(k)$ the output error, μ the step-size, while $(\cdot)^H$ and $(\cdot)^*$ are respectively the Hermitian and complex conjugation operator. Within CLMS, the desired signal $d(k)$ can be assumed to be generated by a strictly linear estimation (SLE) model, given by

$$d(k) = \mathbf{h}^{oH}\mathbf{x}(k) + q(k) \tag{3}$$

where $\mathbf{h}^o = [h_1^o, h_2^o, \ldots, h_N^o]^T$ is the optimal system impulse response vector to be estimated and $q(k)$ is zero-mean independent identically distributed noise with variance σ_q^2.

The seminal paper by Horowitz and Senne [2] considers the situation where the pair of the desired and input signals, $\{d(k), \mathbf{x}(k)\}$, is jointly circularly distributed Gaussian. The zero-mean Gaussian input $\mathbf{x}(k)$ is also implicitly assumed to be second order circular (proper), with a rotation invariant probability density function in the complex domain $\mathbb{C}$ and a vanishing complementary covariance matrix $\mathbf{P} = E[\mathbf{x}(k)\mathbf{x}^T(k)] = \mathbf{0}$ [3], [4]. The work in [2] then provides convergence analysis, in the SLE setting, based on the mean square error (MSE) and the covariance matrix of the weight vector of CLMS. Fisher and Bershad extended the analysis by establishing that the matrix which diagonalises the input covariance matrix also diagonalises the weight error covariance matrix [5]. This makes it possible to analyse the covariance matrix of the weight error vector, which, together with MSE, represents the standard mean square convergence analysis of complex-valued adaptive filtering algorithms [5]. The above assumption of circularity simplifies the analysis framework in many aspects, since in this way complex-valued random signals statistically behave like real-valued ones, and hence, it has been implicitly or explicitly adopted for the analysis of CLMS and its variants in different applications [6]–[13].

Recent advances in so-called *augmented complex statistics* have established a framework for the analysis of second order noncircular random signals, jointly characterised by both the standard covariance matrix $\mathbf{R}$ and complementary covariance matrix $\mathbf{P}$. The non-vanishing complementary covariance may be caused by different powers in the real and imaginary parts or a degree of correlation between the real and imaginary parts, or both. Therefore, in order to make use of all the available second order information, we need to consider the complementary covariance matrix $\mathbf{P}$, in addition to the standard covariance matrix

Manuscript received December 7, 2016; revised May 7, 2017; accepted July 18, 2017. Date of publication August 11, 2017; date of current version August 31, 2017. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Masahiro Yukawa. This work was supported in part by the National Natural Science Foundation of China under Grant 61401094, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20140645, in part by the Fundamental Research Funds for the Central Universities under Grant 2242016K41050, and in part by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry of China. *(Corresponding author: Yili Xia.)*

Y. Xia is with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: yili_xia@seu.edu.cn).

D. P. Mandic is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: d.mandic@imperial.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2017.2739098

$\mathbf{R} = E[\mathbf{x}(k)\mathbf{x}^H(k)]$ [14]–[19]. A concept intimately related to augmented complex statistics is so-called widely linear estimation (WLE), which considers the desired signal, $d_{\mathrm{WL}}(k)$, to be generated by the widely linear model [15], [16], [20]

$$d_{\mathrm{WL}}(k) = \mathbf{h}^{\mathrm{o}H}\mathbf{x}(k) + \mathbf{g}^{\mathrm{o}H}\mathbf{x}^*(k) + q(k) \tag{4}$$

where $\mathbf{g}^{\mathrm{o}} = [g_1^{\mathrm{o}}, g_2^{\mathrm{o}}, \ldots, g_N^{\mathrm{o}}]^T$ is the so-called conjugate optimal system impulse response vector, associated with the input conjugate $\mathbf{x}^*(k)$. The WLE can be therefore considered as a generalised estimation framework in the complex domain $\mathbb{C}$, and has provided modelling advantages over SLE in numerous applications in signal processing, communications, power systems, biomedical engineering and renewable energy [21]–[32]. For instance, the augmented complex statistics have opened the possibility to design LMS-type adaptive algorithms based on the widely linear model, an example of which is the augmented CLMS (ACLMS), given by [24], [25]

$$e(k) = d_{\mathrm{WL}}(k) - \mathbf{h}^H(k)\mathbf{x}(k) - \mathbf{g}^H(k)\mathbf{x}^*(k) \tag{5}$$

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \mu e^*(k)\mathbf{x}(k) \tag{6}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) + \mu e^*(k)\mathbf{x}^*(k) \tag{7}$$

Compared with the standard CLMS, the ACLMS updates an additional weight vector $\mathbf{g}(k)$ in order to track the optimal "conjugate" system impulse response vector, $\mathbf{g}^{\mathrm{o}}$, associated with the desired signal $d_{\mathrm{WL}}(k)$ generated by the widely linear model in (4). Due to the advantages of the widely linear modelling, the ACLMS is second order optimal for both the WLE and SLE of second order circular and noncircular zero-mean Gaussian input signals [32]–[34], since it always yields mean square error (MSE) that is smaller than or equal to that of the strictly linear CLMS.

However, there are several strong motivations, in both theory and practice, for the investigation of the use of CLMS (instead of ACLMS), for the SLE problem in (3) with second order noncircular Gaussian inputs $\mathbf{x}(k)$. On one hand, CLMS remains 'first order' optimal for this task since it gives unbiased estimation of the unknown system weights $\mathbf{h}^{\mathrm{o}}$ in the mean sense [35]. It is even 'second order' optimal in terms of the weight error variance, while ACLMS often encounters a relative performance loss caused by its over-fitting problem, and this performance loss has been proved to increase with an increase in the degree of the noncircularity of the input [36]. Real-world applications of the SLE model include the representation of the received baseband signals over fading channels when the transmission system adopts noncircular complex-valued constellation schemes, e.g., 8 phase shift-keying (8PSK), offset quadrature amplitude modulation (OQAM), and 32QAM [37]–[39]. Its usefulness becomes even more pronounced in multi-user communication systems in the modelling of the receiver in the presence of interference channels where improper Gaussian signals are intentionally employed, because the associated complementary second order statistics provide one more degree of freedom in order to improve the system throughput and reliability [40]–[42]. The SLE model is also adequate to represent the underlying time-series relationships between consecutive complex-valued system voltage samples, which are second order noncircular in unbalanced power systems [43].

The standard mean square convergence behaviour of CLMS for SLE of second order noncircular zero-mean Gaussian inputs is different from that for circular ones in the sense that the nonzero input complementary covariance matrix $\mathbf{P} = E[\mathbf{x}(k)\mathbf{x}^T(k)]$, a counterpart to its standard covariance matrix $\mathbf{R} = E[\mathbf{x}(k)\mathbf{x}^H(k)]$, appears in the evolution of the weight error covariance matrix, and consequently, in the mean square error (MSE) analysis of CLMS [35], [36]. By considering that a joint diagonalisation of both $\mathbf{R}$ and $\mathbf{P}$ with a single singular value decomposition (SVD) is a prerequisite to achieve the closed-form expressions of the weight error covariance vector and the MSE at the steady state, the strong uncorrelating transform (SUT) was introduced [35], [36], with a limitation that a single SVD by SUT is allowed only for a special type of correlated second order noncircular signals. To investigate the steady state mean square performance of CLMS for the general second order noncircular Gaussian signals, the recently introduced approximate uncorrelating transform (AUT) [44] was employed in [34] for a joint single SVD of both $\mathbf{R}$ and $\mathbf{P}$, based on a reasonable approximation. The above analyses directly inherit the principles from the corresponding real domain analyses [45]–[48], and investigate the variances of the output error and the weight error vector, as well as their transient and steady state behaviours. We argue that this reflects only one aspect of the full second order statistics in $\mathbb{C}$, since the complementary second order statistics of the output error and the weight error vector, which reflect their degrees of noncircularity, are not taken into consideration.

Therefore, in this paper, we first propose a novel complementary mean square performance analysis of CLMS for SLE of general second order noncircular zero-mean Gaussian inputs, by investigating the complementary mean square error (CMSE) and the complementary covariance matrix of the weight error vector. We illustrate, for the first time, the evolution of the complementary output error and the weight error vector of CLMS, in terms of their second order noncircularity, when the system input $\mathbf{x}(k)$ or/and the system noise $q(k)$ within the SLE model is/are second order noncircular. The statistical convergence behaviours of CMSE and complementary covariance matrix of the weight error vector are subsequently investigated, and a unified conservative bound on the step-size is derived to guarantee the convergence and stability of both the conventional MSE and the proposed CMSE, for which a joint direct diagonalisation of both the input covariance matrix and its complementary counterpart is enabled by AUT. Next, the closed-form expression for the steady-state complementary MSE (CMSE) of CLMS is obtained, and its dependence on the degree of input noncircularity is quantified for doubly white[1] Gaussian input data. In this way, the proposed complementary mean square analysis augments the standard mean square convergence analysis in [34], and benefiting from the augmented complex statistics, they together

[1]A complex-valued random signal is called doubly white if the off-diagonal elements of its covariance matrix $\mathbf{R}$ and complementary covariance matrix $\mathbf{P}$ are all zeros [16], [18], [19]. Doubly white signals can be both circular and noncircular.

provide the full second order statistical description of CLMS in both transient and steady state stages. The so established full mean square analysis, via a joint consideration of both the standard and complementary mean square analyses, equips us with an additional insight into the mean square evolutions of the real part and imaginary part of the output error and the weight error vector independently, an important property that cannot be achieved by using the standard mean square analysis only. Numerical simulations in the system identification setting support the analysis.

II. Full Mean Square Analysis of CLMS

Consider the use of CLMS, summarised in (1) and (2), for SLE of a desired response $d(k)$, given in (3). Upon introducing the $N \times 1$ weight error vector

$$\widetilde{\mathbf{h}}(k) = \mathbf{h}(k) - \mathbf{h}^{\mathrm{o}} \tag{8}$$

the output error $e(k)$ in (1) becomes

$$e(k) = \ q(k) - \widetilde{\mathbf{h}}^{H}(k)\mathbf{x}(k) \tag{9}$$

while for its conjugate we have

$$e^*(k) = q^*(k) - \widetilde{\mathbf{h}}^{T}(k)\mathbf{x}^*(k) \tag{10}$$

From (2), the recursion for the evolution of the weight error vector $\widetilde{\mathbf{h}}(k)$ now becomes

$$\begin{aligned}\widetilde{\mathbf{h}}(k+1) &= \ \widetilde{\mathbf{h}}(k) + \mu\big(q^*(k) - \widetilde{\mathbf{h}}^{T}(k)\mathbf{x}^*(k)\big)\mathbf{x}(k) \\ &= \ \big(\mathbf{I} - \mu\mathbf{x}(k)\mathbf{x}^{H}(k)\big)\widetilde{\mathbf{h}}(k) + \mu q^*(k)\mathbf{x}(k) \end{aligned} \tag{11}$$

where $\mathbf{I}$ is an $N \times N$ identity matrix. Based on (11), the evolution of $\widetilde{\mathbf{h}}^{H}(k+1)$ and $\widetilde{\mathbf{h}}^{T}(k+1)$ can be recursively expressed as

$$\widetilde{\mathbf{h}}^{H}(k+1) = \widetilde{\mathbf{h}}^{H}(k)\big(\mathbf{I} - \mu\mathbf{x}(k)\mathbf{x}^{H}(k)\big) + \mu q(k)\mathbf{x}^{H}(k) \tag{12}$$

$$\widetilde{\mathbf{h}}^{T}(k+1) = \widetilde{\mathbf{h}}^{T}(k)\big(\mathbf{I} - \mu\mathbf{x}^*(k)\mathbf{x}^{T}(k)\big) + \mu q^*(k)\mathbf{x}^{T}(k) \tag{13}$$

A. Summary of Standard Mean Square Analysis

The standard mean square error (MSE) performance criterion of CLMS, denoted by $J(k)$, can be defined as

$$J(k) = E[|e(k)|^2] = E[e(k)e^*(k)] \tag{14}$$

where $E[\cdot]$ is the statistical expectation operator, and $|\cdot|$ denotes the absolute value of a complex number. In order to aid the theoretical mean square performance analysis of CLMS with mathematical tractability, the standard independence assumptions commonly used in adaptive filtering analyses, that is, the system noise $q(k)$ is statistically independent of any other signal in the CLMS algorithm and the weight error vector $\widetilde{\mathbf{h}}(k)$ is statistically independent of the adaptive filter input $\mathbf{x}(k)$ [49], have been adopted in [2], [5], [6], [10], [34], [35] to decompose $J(k)$ in (14) as

$$\begin{aligned} J(k) &= \ \sigma_q^2 + E[\widetilde{\mathbf{h}}^{H}(k)\mathbf{x}(k)\mathbf{x}^{H}(k)\widetilde{\mathbf{h}}(k)] \\ &= \ \sigma_q^2 + \mathrm{tr}[\mathbf{R}\mathbf{K}(k)] \end{aligned} \tag{15}$$

where $\mathrm{tr}[\cdot]$ is the matrix trace operator, and

$$\mathbf{K}(k) = E[\widetilde{\mathbf{h}}(k)\widetilde{\mathbf{h}}^{H}(k)] \tag{16}$$

is the weight error covariance matrix. However, we should mention that in some practical applications, such as blind channel equalisation, the independence between the $\widetilde{\mathbf{h}}(k)$ and $\mathbf{x}(k)$ may be violated [50]. The mean square performance analysis of CLMS in (15) now rests upon the standard second order characteristics of the weight error vector $\widetilde{\mathbf{h}}(k)$. By multiplying both sides of (11) and (12), taking the statistical expectation $E[\cdot]$, and using the standard independence assumptions, the evolution of the weight error covariance matrix $\mathbf{K}(k)$ in (16) is obtained in the form [34], [35]

$$\begin{aligned}\mathbf{K}(k+1) &= \mathbf{K}(k) - \mu\big(\mathbf{R}\mathbf{K}(k) + \mathbf{K}(k)\mathbf{R}\big) \\ &\quad + \mu^2\big(\sigma_q^2\mathbf{R} + E[\mathbf{x}(k)\mathbf{x}^{H}(k)\widetilde{\mathbf{h}}(k)\widetilde{\mathbf{h}}^{H}(k)\mathbf{x}(k)\mathbf{x}^{H}d(k)]\big)\end{aligned} \tag{17}$$

Under the independence assumptions, the (i,j)th entry of the expectation matrix of the last term on the right hand side (RHS) of (17) becomes

$$\begin{aligned}&\big\{E[\mathbf{x}(k)\mathbf{x}^{H}(k)\widetilde{\mathbf{h}}(k)\widetilde{\mathbf{h}}^{H}(k)\mathbf{x}(k)\mathbf{x}^{H}(k)]\big\}_{ij} \\ &\quad = \sum_{l=1}^{N}\sum_{m=1}^{N} E[x_i(k)x_l^*(k)x_m(k)x_j^*(k)]E[\widetilde{h}_l(k)\widetilde{h}_m^*(k)]\end{aligned}$$

Upon employing the Gaussian fourth order moment factorising theorem, we have

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

and hence [34], [35],

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

Thus, the evolution of the weight error covariance matrix $\mathbf{K}(k)$ in (17) now becomes [34], [35]

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

Equations (15) and (18) together describe the standard mean square performance of the CLMS algorithm for second order noncircular Gaussian input data, and its dependence on the input noncircularity, represented by the complementary covariance matrix $\mathbf{P}$, can be clearly observed.

B. Proposed Complementary Mean Square Analysis

Since the zero-mean Gaussian input $\mathbf{x}(k)$ and the system noise $q(k)$ are considered to be second order noncircular, it is also natural to investigate whether the so introduced noncircularity propagates into the error $e(k)$ and the weight error vector $\widetilde{\mathbf{h}}(k)$, the two key system parameters which govern the CLMS algorithm[2]. By virtue of the augmented complex statistics

[2]This information will give key new insights into the behaviour of CLMS for general inputs, and such analysis has not been conducted so far.

[14]–[19], we can now first define the complementary mean square error (CMSE) $\widetilde{J}(k)$ as

$$\widetilde{J}(k) = E[e^2(k)] \tag{19}$$

to represent the noncircularity of the output error $e(k)$. Based on (9) and using the standard assumptions discussed above, $\widetilde{J}(k)$ can be further evaluated as

$$\begin{aligned}\widetilde{J}(k) =& \; E[(q(k) - \widetilde{\mathbf{h}}^H(k)\mathbf{x}(k))^2] \\ =& \; E[q^2(k)] + E[\widetilde{\mathbf{h}}^H(k)\mathbf{x}(k)\mathbf{x}^T(k)\widetilde{\mathbf{h}}^*(k)] \\ =& \; \widetilde{\sigma}_q^2 + \mathrm{tr}[\mathbf{P}\widetilde{\mathbf{K}}^*(k)]\end{aligned} \tag{20}$$

in which $\widetilde{\sigma}_q^2 = E[q^2(k)]$ is the complementary variance (pseudo-variance) of the second order noncircular system noise $q(k)$, and

$$\widetilde{\mathbf{K}}(k) = E[\widetilde{\mathbf{h}}(k)\widetilde{\mathbf{h}}^T(k)] \tag{21}$$

is the complementary covariance matrix of the weight error vector $\widetilde{\mathbf{h}}(k)$. Upon multiplying both sides of (11) by $\widetilde{\mathbf{h}}^T(k+1)$ in (13), and again applying the statistical expectation operator and the standard independence assumptions, we obtain

$$\begin{aligned}\widetilde{\mathbf{K}}(k+1) =& \; \widetilde{\mathbf{K}}(k) - \mu\big(\mathbf{R}\widetilde{\mathbf{K}}(k) + \widetilde{\mathbf{K}}(k)\mathbf{R}^*\big) \\ &+ \mu^2\big(\widetilde{\sigma}_q^{2*}\mathbf{P} + E[\mathbf{x}(k)\mathbf{x}^H(k)\widetilde{\mathbf{h}}(k)\widetilde{\mathbf{h}}^T(k)\mathbf{x}^*(k)\mathbf{x}^T(k)]\big)\end{aligned} \tag{22}$$

The (i,j)th entry of the statistical expectation of the last term on the RHS of (22) can be expressed as

$$\begin{aligned}&\big\{E[\mathbf{x}(k)\mathbf{x}^H(k)\widetilde{\mathbf{h}}(k)\widetilde{\mathbf{h}}^T(k)\mathbf{x}^*(k)\mathbf{x}^T(k)]\big\}_{ij} \\ =& \sum_{l=1}^{N}\sum_{m=1}^{N} E[x_i(k)x_l^*(k)x_m^*(k)x_j(k)]E[\widetilde{h}_l(k)\widetilde{h}_m(k)]\end{aligned}$$

Upon employing the Gaussian fourth order moment factorising theorem, we now have

$$E[x_i(k)x_l^*(k)x_m^*(k)x_j(k)] = r_{il}r_{jm} + r_{im}r_{jl} + p_{ij}p_{lm}^*$$

and hence

$$\begin{aligned}&E\;[\mathbf{x}(k)\mathbf{x}^H(k)\widetilde{\mathbf{h}}(k)\widetilde{\mathbf{h}}^T(k)\mathbf{x}^*(k)\mathbf{x}^T(k)] \\ =& \; 2\mathbf{R}\widetilde{\mathbf{K}}(k)\mathbf{R}^* + \mathbf{P}\mathrm{tr}[\mathbf{P}^*\widetilde{\mathbf{K}}(k)]\end{aligned}$$

Therefore, the recursion for the complementary covariance matrix of the weight error vector $\widetilde{\mathbf{h}}(k)$, that is, $\widetilde{\mathbf{K}}(k)$ in (21), now becomes

$$\begin{aligned}\widetilde{\mathbf{K}}(k+1) =& \; \widetilde{\mathbf{K}}(k) - \mu\big(\mathbf{R}\widetilde{\mathbf{K}}(k) + \widetilde{\mathbf{K}}(k)\mathbf{R}^*\big) \\ &+ \mu^2\big(\widetilde{\sigma}_q^{2*}\mathbf{P} + 2\mathbf{R}\widetilde{\mathbf{K}}(k)\mathbf{R}^* + \mathbf{P}\mathrm{tr}[\mathbf{P}^*\widetilde{\mathbf{K}}(k)]\big)\end{aligned} \tag{23}$$

Remark 1: Equations (20) and (23) indicate that both $e(k)$ and $\widetilde{\mathbf{h}}(k)$ can be second order noncircular if the Gaussian input $\mathbf{x}(k)$ or the system noise $q(k)$ is second order noncircular. Both equations describe the complementary mean square performance of CLMS, and together with the standard mean square performance provided in (15) and (18), they provide a complete description on the second order statistical behaviour of CLMS, based on the augmented complex statistics [14]–[19].

III. Full Mean Square Stability of CLMS

We now consider sufficient conditions for the full mean square convergence of the weight error vector $\widetilde{\mathbf{h}}(k)$ of CLMS for the general second order noncircular Gaussian input data. This is challenging in the sense that it requires a simultaneous diagonalisation of both the covariance matrix $\mathbf{R}$ and the complementary covariance matrix $\mathbf{P}$ in (18) and (23). The first attempt in this direction was the work in [35], which applies the strong uncorrelating transform (SUT) [51], [52] on $\mathbf{R}$ and $\mathbf{P}$. However, since SUT admits a single SVD for both matrices only for a special type of correlated second order noncircular signals, the analysis in [35] cannot be straightforwardly extended to the case of general second order noncircular signals for which the off-diagonal elements in both $\mathbf{R}$ and $\mathbf{P}$ contain nonzero elements. In order to both address the diagonalisation problem encountered by SUT and simplify the analysis, we here employ the recently introduced approximate uncorrelating transform (AUT), which allows for the diagonalisation of $\mathbf{R}$ and $\mathbf{P}$ based on a single singular value decomposition (SVD), within some reasonable approximations [34], [44].

The Takagi factorisation states that any complex-valued symmetric matrix, like the complementary covariance matrix $\mathbf{P} = \mathbf{P}^T$, can be diagonalised as

$$\mathbf{P} = \mathbf{Q}\mathbf{\Lambda}_\mathbf{p}\mathbf{Q}^T \tag{24}$$

while for its conjugate

$$\mathbf{P}^* = \mathbf{Q}^*\mathbf{\Lambda}_\mathbf{p}\mathbf{Q}^H \tag{25}$$

where $\mathbf{Q}$ is a unitary matrix, $\mathbf{Q}\mathbf{Q}^H = \mathbf{I}$, and $\mathbf{\Lambda}_\mathbf{p} = \mathrm{diag}\{p_1, p_2, \ldots, p_N\}$ is a diagonal matrix of real-valued entries, where $p_1 \geq p_2 \geq \cdots \geq p_N$ are the nonnegative square roots of $\mathbf{P}\mathbf{P}^H$ [53].

The approximate uncorrelating transform (AUT) [44] states that this same matrix $\mathbf{Q}$ can be used to approximately diagonalise the covariance matrix $\mathbf{R}$, so that

$$\mathbf{R} \simeq \mathbf{Q}\mathbf{\Lambda}_\mathbf{r}\mathbf{Q}^H \tag{26}$$

and hence its inversion

$$\mathbf{R}^{-1} \simeq \mathbf{Q}\mathbf{\Lambda}_\mathbf{r}^{-1}\mathbf{Q}^H \tag{27}$$

where $\mathbf{\Lambda}_\mathbf{r} = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_N\}$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$, and λ_i are the real-valued eigenvalues of $\mathbf{R}$. The approximations in (26) and (27) are valid for univariate data, and the equality is achieved when $\mathbf{x}(k)$ is real-valued, that is, maximum noncircular [34], [44]. The benefits of using AUT to obtain the stability bound on the step-size μ in the mean square sense, and for the closed-form solutions for the steady state performance of CLMS, are illustrated in the following analysis.

With the use of AUT, we can now rotate the weight error vector and the input vector to give

$$\hat{\mathbf{h}}(k) = \mathbf{Q}^H\widetilde{\mathbf{h}}(k) \;\text{ and }\; \hat{\mathbf{x}}(k) = \mathbf{Q}^H\mathbf{x}(k) \tag{28}$$

and to decompose the term $\mathrm{tr}[\mathbf{R}\mathbf{K}(k)]$ on the RHS of (18) as

$$\begin{aligned}\mathrm{tr}[\mathbf{R}\mathbf{K}(k)] &= E[\widetilde{\mathbf{h}}^H(k)\mathbf{x}(k)\mathbf{x}^H(k)\widetilde{\mathbf{h}}(k)] \\ &= E[\hat{\mathbf{h}}^H(k)\boldsymbol{\Lambda}_{\mathbf{r}}\hat{\mathbf{h}}(k)] \\ &= \boldsymbol{\lambda}_{\mathbf{r}}^T\boldsymbol{\kappa}(k) \end{aligned} \tag{29}$$

where $\boldsymbol{\lambda}_{\mathbf{r}} = [\lambda_1, \lambda_2, \ldots, \lambda_N]^T$, and $\boldsymbol{\kappa}(k)$ is the $N \times 1$ weight error variance vector after data diagonalisation, the components of which are the diagonal elements of $E[\hat{\mathbf{h}}(k)\hat{\mathbf{h}}^H(k)]$, defined as

$$\boldsymbol{\kappa}(k) = \left[E[|\hat{h}_1(k)|^2], E[|\hat{h}_2(k)|^2], \ldots, E[|\hat{h}_N(k)|^2]\right]^T \tag{30}$$

Then, based on (18), we obtain the evolution of $\boldsymbol{\kappa}(k)$ for general second order noncircular inputs, as discussed in [34], [35], given by

$$\begin{aligned}\boldsymbol{\kappa}(k+1) = &\underbrace{\left(\mathbf{I} - 2\mu\boldsymbol{\Lambda}_{\mathbf{r}} + \mu^2\left(\boldsymbol{\Lambda}_{\mathbf{r}}^2 + \boldsymbol{\Lambda}_{\mathbf{p}}^2 + \boldsymbol{\lambda}_{\mathbf{r}}\boldsymbol{\lambda}_{\mathbf{r}}^T\right)\right)}_{\mathbf{F}}\boldsymbol{\kappa}(k) \\ &+ \mu^2\sigma_q^2\boldsymbol{\lambda}_{\mathbf{r}}\end{aligned} \tag{31}$$

We shall now consider the diagonalisation of the complementary weight error covariance matrix $\widetilde{\mathbf{K}}(k)$. To this end, we first observe that by using (25) and (28), the term $\mathrm{tr}[\mathbf{P}^*\widetilde{\mathbf{K}}(k)]$ on the RHS of (23) can be decomposed as

$$\begin{aligned}\mathrm{tr}[\mathbf{P}^*\widetilde{\mathbf{K}}(k)] &= E[\widetilde{\mathbf{h}}^T(k)\mathbf{x}^*(k)\mathbf{x}^H(k)\widetilde{\mathbf{h}}(k)] \\ &= E[\hat{\mathbf{h}}^T(k)\boldsymbol{\Lambda}_{\mathbf{p}}\hat{\mathbf{h}}(k)] \\ &= \boldsymbol{\lambda}_{\mathbf{p}}^T\widetilde{\boldsymbol{\kappa}}(k) \end{aligned} \tag{32}$$

where $\boldsymbol{\lambda}_{\mathbf{p}} = [p_1, p_2, \ldots, p_N]^T$ and $\widetilde{\boldsymbol{\kappa}}(k)$ is the complementary weight error variance vector, the components of which are the diagonal elements of $E[\hat{\mathbf{h}}(k)\hat{\mathbf{h}}^T(k)]$, defined as

$$\widetilde{\boldsymbol{\kappa}}(k) = \left[E[\hat{h}_1^2(k)], E[\hat{h}_2^2(k)], \ldots, E[\hat{h}_N^2(k)]\right]^T \tag{33}$$

Based on (23) and (32), and after a few algebraic manipulations, the evolution of $\widetilde{\boldsymbol{\kappa}}(k)$ can be expressed as

$$\widetilde{\boldsymbol{\kappa}}(k+1) = \underbrace{\left(\mathbf{I} - 2\mu\boldsymbol{\Lambda}_{\mathbf{r}} + \mu^2\left(2\boldsymbol{\Lambda}_{\mathbf{r}}^2 + \boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T\right)\right)}_{\widetilde{\mathbf{F}}}\widetilde{\boldsymbol{\kappa}}(k) + \mu^2\widetilde{\sigma}_q^{2*}\boldsymbol{\lambda}_{\mathbf{p}} \tag{34}$$

Remark 2: From (34), observe that when the Gaussian input $\mathbf{x}(k)$ and the system noise $q(k)$ are second order noncircular, the driving term $\mu^2\widetilde{\sigma}_q^{2*}\boldsymbol{\lambda}_{\mathbf{p}}$ for the recursion of $\widetilde{\boldsymbol{\kappa}}(k)$ in (34) does exist. Their noncircularity therefore propagates into both the weight error vector, $\widetilde{\mathbf{h}}(k)$ and the output error, $e(k)$, the noncircularity of which is measured by the complementary cost $\widetilde{J}(k)$, according to (32) and (20). The noncircularity of $\widetilde{\mathbf{h}}(k)$ vanishes when either $\mathbf{x}(k)$ or $q(k)$ is second order circular, since $\widetilde{\boldsymbol{\kappa}}(k)$ would then converge to $\mathbf{0}$. While the noncircularity of $e(k)$ vanishes when both $\mathbf{x}(k)$ and $q(k)$ are second order circular, notice that its noncircularity degree would be a constant governed by the complementary variance $\widetilde{\sigma}_q^2$, only when $q(k)$ is second order noncircular.

Strictly speaking, we are not able to talk about the convergence for a complex-valued vector, like $\widetilde{\boldsymbol{\kappa}}(k)$ in (34), since the complex domain $\mathbb{C}$ is not ordered. However, the real-valued nature of $\widetilde{\mathbf{F}}$ guarantees that the real and imaginary parts of $\widetilde{\boldsymbol{\kappa}}(k)$ evolve independently, except for the shared transition matrix $\widetilde{\mathbf{F}}$, and that the convergence of $\widetilde{\boldsymbol{\kappa}}(k)$ implies that both its real and imaginary parts converge, which would consequently lead to the convergence of the CMSE $\widetilde{J}(k)$ in the same way, according to (20), (32) and (34). Based on (31) and (34), the full mean square convergence of CLMS is then subject to the condition that all the eigenvalues of the transition matrices $\mathbf{F}$ and $\widetilde{\mathbf{F}}$ are less than unity.

We shall first consider the upper bound on μ in order to establish the stability of $\widetilde{\mathbf{F}}$. To achieve this, we here follow the classic work of A. Feuer and E. Weinstein in [45], and consider that the eigenvalues of $\widetilde{\mathbf{F}}$ in (34) are the solutions of an equation in α, so that

$$\det[\widetilde{\mathbf{F}} - \alpha\mathbf{I}] = 0 \tag{35}$$

where $\det[\cdot]$ is the matrix determinant operator. By defining

$$\rho_i = 1 - 2\mu\lambda_i + 2\mu^2\lambda_i^2,\ i = 1, 2, \ldots, N \tag{36}$$

the matrix $\widetilde{\mathbf{F}}$ now can be expressed as

$$\widetilde{\mathbf{F}} = \mathrm{diag}\{\rho_1, \rho_2, \ldots, \rho_N\} + \mu^2\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T$$

and hence

$$\begin{aligned}&\det\left[\widetilde{\mathbf{F}} - \alpha\mathbf{I}\right] \\ &\quad= \det\left[\mathrm{diag}\{\rho_1 - \alpha, \rho_2 - \alpha, \ldots, \rho_N - \alpha\} + \mu^2\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T\right] \\ &\quad= \det\left[\mathrm{diag}\{\rho_1 - \alpha, \rho_2 - \alpha, \ldots, \rho_N - \alpha\}\right] \\ &\qquad\cdot \det\left[\mathbf{I} + \mu^2\boldsymbol{\lambda}_{\mathbf{p}}^T\mathrm{diag}\left\{\frac{1}{\rho_1 - \alpha}, \frac{1}{\rho_2 - \alpha}, \ldots, \frac{1}{\rho_N - \alpha}\right\}\boldsymbol{\lambda}_{\mathbf{p}}\right] \\ &\quad= \left[\prod_{i=1}^N(\rho_i - \alpha)\right]\cdot\left[1 + \sum_{i=1}^N\frac{\mu^2p_i^2}{\rho_i - \alpha}\right] = 0\end{aligned}$$

Since the common denominator of the sum of terms $\sum_{i=1}^N p_i^2/(\rho_i - \alpha)$ is $\prod_{i=1}^N(\rho_i - \alpha)$, we only need to consider the equation

$$f(\alpha) = 1 + \sum_{i=1}^N\frac{\mu^2p_i^2}{\rho_i - \alpha} = 0 \tag{37}$$

Clearly, the poles of $f(\alpha)$ are ρ_i, and from (36)

$$\rho_i = 1 - 2\mu\lambda_i + 2\mu^2\lambda_i^2 = (1 - \mu\lambda_i)^2 + \mu^2\lambda_i^2 > 0 \tag{38}$$

Therefore, the poles of $f(\alpha)$ are placed on the positive real axis. From (37), we further observe that

$$\frac{\partial f(\alpha)}{\partial\alpha} = \sum_{i=1}^N\frac{\mu^2p_i^2}{(\rho_i - \alpha)^2} > 0 \tag{39}$$

which indicates that $f(\alpha)$ must therefore assume a strict form in the sense that between each pair of successive poles, the function $f(\alpha)$ has a single zero, as illustrated in Fig. 1. If we arrange the poles in an increasing order, i.e., $0 < \rho_1 < \rho_2 < \cdots < \rho_N$, then

$$0 < \rho_1 < \alpha_1 < \rho_2 < \alpha_2 < \cdots < \alpha_{N-1} < \rho_N < \alpha_N \tag{40}$$

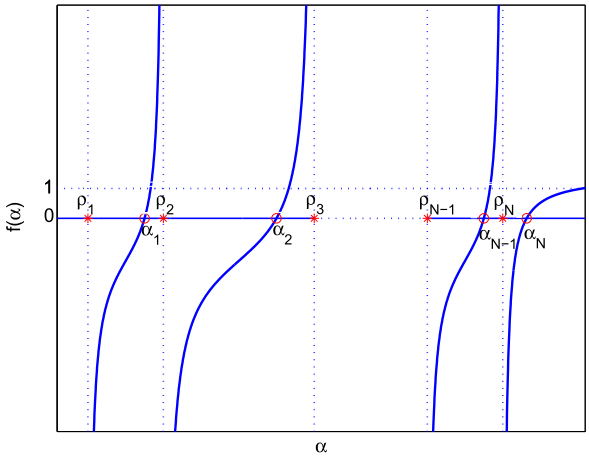


Fig. 1. Geometric illustration of the general form of $f(\alpha)$.

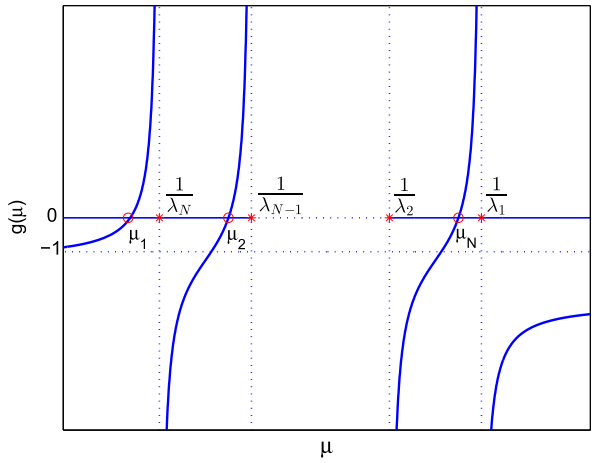


Fig. 2. Geometric illustration of the general form of $g(\mu)$.

where α_i are the zeros of $f(\alpha)$, and hence the eigenvalues of $\widetilde{\mathbf{F}}$. The stability requirement $\alpha_N < 1$ means that all the eigenvalues of $\widetilde{\mathbf{F}}$ are located within the unit circle, and hence the recursion in (34) is stable. From Fig. 1, this is the case if and only if

$$\rho_i < 1, \ i = 1, 2, \ldots, N \tag{41}$$

and

$$f(\alpha) \mid_{\alpha=1} = 1 + \sum_{i=1}^{N} \frac{\mu^2 p_i^2}{\rho_i - \alpha} > 0 \tag{42}$$

Substituting (36) into (41) and (42) yields the following necessary and sufficient conditions

$$0 < \mu < \frac{1}{\lambda_i}, \ i = 1, 2, \ldots, N \tag{43}$$

and

$$\sum_{i=1}^{N} \frac{\mu p_i^2}{2(\lambda_i - \mu\lambda_i^2)} - 1 < 0 \tag{44}$$

It is of great practical interest to translate the conditions in (43) and (44) into direct bounds on the step-size μ. To this end, we first denote the left hand side of (44) by $g(\mu)$, to give

$$g(\mu) = \sum_{i=1}^{N} \frac{\mu p_i^2}{2(\lambda_i - \mu\lambda_i^2)} - 1 \tag{45}$$

This is a monotonically increasing function of μ since

$$\frac{\partial g(\mu)}{\partial \mu} = \sum_{i=1}^{N} \frac{2p_i^2 \lambda_i}{4(\lambda_i - \mu\lambda_i^2)^2} > 0 \tag{46}$$

Moreover, $g(\mu)$ has poles at $1/\lambda_i, i = 1, 2, \ldots, N$, all of which are located on the positive real axis, its values are equal to -1 for $\mu = 0$, and $\lim_{\mu\to\infty} g(\mu) < 0$. Therefore, if $\mu_i, i = 1, 2, \ldots, N$, are the roots of $g(\mu) = 0$, and are arranged in an increasing order, i.e., $0 < \mu_1 < \mu_2 < \cdots < \mu_N$, then $g(\mu)$ strictly obeys the form that between each pair of successive zeros, $g(\mu)$ has a single pole, as illustrated in Fig. 2, and hence

$$0 < \mu_1 < \frac{1}{\lambda_N} < \mu_2 < \frac{1}{\lambda_{N-1}} < \cdots < \mu_N < \frac{1}{\lambda_1} \tag{47}$$

It is now clear that the condition $0 < \mu < \mu_1$ is equivalent to conditions in (43) and (44). In general, a closed-form solution for μ cannot be obtained. However, we will take the advantage of a theorem established in [54] to obtain a tight lower bound on μ_1. To achieve this, first note that after a few mathematical manipulations, the polynomial $g(\mu) = 0$ can be rewritten in the form

$$\left(\frac{1}{\mu}\right)^N - l_1 \left(\frac{1}{\mu}\right)^{N-1} + l_2 \left(\frac{1}{\mu}\right)^{N-2} + \cdots + (-1)^N l_N = \prod_{i=1}^{N} \left(\frac{1}{\mu} - \frac{1}{\mu_i}\right) = 0 \tag{48}$$

where

$$l_1 = \frac{1}{2} \sum_{i=1}^{N} \left(2\lambda_i + \frac{p_i^2}{\lambda_i}\right) \tag{49}$$

$$l_2 = \sum_{i=1}^{N} \sum_{j>i}^{N} \left(\lambda_i + \frac{p_i^2}{\lambda_i}\right) \lambda_j \tag{50}$$

In this setting, μ_1 is the smallest root of (48), and the theorem developed in [54] shows that μ_1 is lower bounded by

$$\mu_1 \geq \frac{N}{l_1 + \sqrt{(N-1)^2 l_1^2 - 2N(N-1) l_2}} = \mu_0 \tag{51}$$

Thus, to ensure that $\widetilde{\boldsymbol{\kappa}}(k)$ in (34) converges in the mean square sense, μ should be bounded by

$$0 < \mu < \mu_0 \tag{52}$$

To make (52) more practical, from (51) we note that

$$\mu_0 \geq \frac{1}{l_1} = \frac{2}{\sum_{i=1}^{N} \left(2\lambda_i + \frac{p_i^2}{\lambda_i}\right)} \tag{53}$$

Then

$$0 < \mu < \frac{1}{\sum_{i=1}^{N}\left(\lambda_i + \frac{p_i^2}{2\lambda_i}\right)} \tag{54}$$

Using the same procedure to analyse the transition matrix $\mathbf{F}$, as shown in Appendix A, we can find that the conservative bound on μ which enables the mean-square convergence of $\boldsymbol{\kappa}(k)$ in exactly the same way as that given in (54) for $\widetilde{\boldsymbol{\kappa}}(k)$.

Remark 3: Equation (54) provides a unified conservative bound on μ for the full mean square convergence of CLMS for second order noncircular Gaussian inputs. It indicates that the step-size μ within this bound, which enables the recursive minimisation of the MSE $J(k) = E[|e(k)|^2]$, also guarantees the convergence of the second order noncircularity of the output error $e(k)$, represented by the CMSE $\widetilde{J}(k) = E[e^2(k)]$, in terms of uncoupled evolutions of its real and imaginary components. It is also interesting to observe that when $\mathbf{x}(k)$ is real-valued, that is, maximum noncircular, the upper bound in (54) becomes $2/(3\mathrm{tr}[\mathbf{R}])$, identical to that derived for mean square convergence of real-valued LMS [45], [55]. This is because for a real-valued $\mathbf{x}(k)$, we have $\mathbf{R} = \mathbf{P} = \mathbf{P}^*$ and $p_i = \lambda_i$, and so the transition matrix $\mathbf{F}$ coincides with $\widetilde{\mathbf{F}}$, and is also identical to that of the real-valued LMS algorithm. Note that 2μ is used in the LMS recursion in [45], which yields a slightly different expression for the transition matrix used in [45]. When $\mathbf{x}(k)$ is second order circular, we have $p_i = 0$, and the upper bound in (54) becomes $1/\mathrm{tr}[\mathbf{R}]$, which, as desired, coincides with the existing results in [2], [10], [11].

IV. Full Mean Square Steady State Analysis of CLMS

Suppose that μ is chosen such that the full mean square stability of CLMS is guaranteed. To perform a full steady state performance analysis on CLMS with second order noncircular Gaussian inputs, observe that when $k \to \infty$, then based on (15) and (29) the steady state MSE (SS-MSE) of CLMS, denoted by $J(\infty)$, can be expressed as

$$J(\infty) = \sigma_q^2 + \boldsymbol{\lambda}_{\mathbf{r}}^T \boldsymbol{\kappa}(\infty) \tag{55}$$

In a similar way, based on (20) and (32), the steady state complementary MSE (SS-CMSE) of CLMS, denoted by $\widetilde{J}(\infty)$, can be derived as

$$\widetilde{J}(\infty) = \widetilde{\sigma}_q^2 + \boldsymbol{\lambda}_{\mathbf{p}}^T \widetilde{\boldsymbol{\kappa}}^*(\infty) \tag{56}$$

Based on (34), the steady-state value of $\widetilde{\boldsymbol{\kappa}}(k)$ is given by

$$\widetilde{\boldsymbol{\kappa}}(\infty) = \left(2\boldsymbol{\Lambda}_{\mathbf{r}} - 2\mu\boldsymbol{\Lambda}_{\mathbf{r}}^2 - \mu\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T\right)^{-1} \mu\widetilde{\sigma}_q^{2*}\boldsymbol{\lambda}_{\mathbf{p}} \tag{57}$$

Upon defining

$$\boldsymbol{\Lambda}_{\mathbf{1}} = 2\boldsymbol{\Lambda}_{\mathbf{r}} - 2\mu\boldsymbol{\Lambda}_{\mathbf{r}}^2 \tag{58}$$

and employing the matrix inversion lemma, we have

$$\begin{aligned}
&\left(\boldsymbol{\Lambda}_{\mathbf{1}} - \mu\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T\right)^{-1} \\
&\quad = \left(\boldsymbol{\Lambda}_{\mathbf{1}}(\mathbf{I} - \mu\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T)\right)^{-1} \\
&\quad = \left(\mathbf{I} + (\mathbf{I} - \mu\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T)^{-1}\mu\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T\right)\boldsymbol{\Lambda}_{\mathbf{1}}^{-1} \\
&\quad = \left(\mathbf{I} + \mu\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}(\mathbf{I} - \boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T\mu\boldsymbol{\Lambda}_{\mathbf{1}}^{-1})^{-1}\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T\right)\boldsymbol{\Lambda}_{\mathbf{1}}^{-1} \\
&\quad = \left(\mathbf{I} + \mu\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}(1 - \boldsymbol{\lambda}_{\mathbf{p}}^T\mu\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}})^{-1}\boldsymbol{\lambda}_{\mathbf{p}}^T\right)\boldsymbol{\Lambda}_{\mathbf{1}}^{-1} \\
&\quad = \left(\mathbf{I} + \frac{\mu\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T}{1 - \mu\boldsymbol{\lambda}_{\mathbf{p}}^T\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}}\right)\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}
\end{aligned} \tag{59}$$

Upon substituting (59) into (57), we have

$$\widetilde{\boldsymbol{\kappa}}(\infty) = \mu\widetilde{\sigma}_q^{2*}\left(\mathbf{I} + \frac{\mu\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T}{1 - \mu\boldsymbol{\lambda}_{\mathbf{p}}^T\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}}\right)\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}$$

while its conjugate takes the form

$$\widetilde{\boldsymbol{\kappa}}^*(\infty) = \mu\widetilde{\sigma}_q^2\left(\mathbf{I} + \frac{\mu\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}\boldsymbol{\lambda}_{\mathbf{p}}^T}{1 - \mu\boldsymbol{\lambda}_{\mathbf{p}}^T\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}}}\right)\boldsymbol{\Lambda}_{\mathbf{1}}^{-1}\boldsymbol{\lambda}_{\mathbf{p}} \tag{60}$$

Upon substituting (60) into (56), and after a few algebraic manipulations, we obtain the closed-form expression of the steady state complementary MSE (SS-CMSE), $\widetilde{J}(\infty)$, as

$$\widetilde{J}(\infty) = \widetilde{\sigma}_q^2 + \frac{\sum_{i=1}^{N}\frac{\mu\widetilde{\sigma}_q^2 p_i^2}{2\lambda_i - 2\mu\lambda_i^2}}{1 - \sum_{i=1}^{N}\frac{\mu p_i^2}{2\lambda_i - 2\mu\lambda_i^2}} \tag{61}$$

In a similar way, the closed-form expression of the standard SS-MSE of CLMS, $J(\infty)$ in (55), can be obtained as [34]

$$J(\infty) = \sigma_q^2 + \frac{\sum_{i=1}^{N}\frac{\mu\sigma_q^2\lambda_i^2}{2\lambda_i - \mu\lambda_i^2 - \mu p_i^2}}{1 - \sum_{i=1}^{N}\frac{\mu\lambda_i^2}{2\lambda_i - \mu\lambda_i^2 - \mu p_i^2}} \tag{62}$$

Remark 4: Equations (61) and (62) jointly provide a complete view on the steady state performance of CLMS for general second order noncircular Gaussian inputs, whereby the off-diagonal elements of both $\mathbf{R}$ and $\mathbf{P}$ contain nonzero elements. The full steady state performance of CLMS is governed by the augmented second order statistics of input data (both covariance and complementary covariance), in terms of λ_i and p_i, and by system noise, in terms of its variance σ_q^2, complementary variance $\widetilde{\sigma}_q^2$, and the step-size μ.

A. Steady State Performance of CLMS vs. Degree of Input Noncircularity

It is of particular interest to find an explicit link between the degree of input noncircularity and the full steady state performance of CLMS. To simplify our analysis, we consider a setting where the second order noncircular Gaussian input $\mathbf{x}(k)$

is doubly white (DW), for which $\mathbf{x}_r(k) \perp \mathbf{x}_\imath(k)$ and

$$\mathbf{R} = \sigma_x^2\mathbf{I} \text{ and } \mathbf{P} = \widetilde{\sigma}_x^2\mathbf{I} \tag{63}$$

where $\sigma_x^2 = E[x(k)x^*(k)]$ and $\widetilde{\sigma}_x^2 = E[x^2(k)]$ are respectively the variance and complementary variance of $\mathbf{x}(k)$ [18], [19]. Since the diagonal elements $p_i, i = 1, 2, \ldots, N$, in $\mathbf{\Lambda}_p$, obtained by the Takagi factorisation, are nonnegative square roots of $\mathbf{P}\mathbf{P}^H$ [53], from (63), for doubly white second order noncircular data, we have $\mathbf{P}\mathbf{P}^H = |\widetilde{\sigma}_x^2|^2\mathbf{I}$, and hence

$$\lambda_i = \sigma_x^2 \text{ and } p_i = |\widetilde{\sigma}_x^2| \tag{64}$$

Consider a measure, η, of the degree of input noncircularity in a signal, defined as the ratio of the absolute value of its complementary covariance $\widetilde{\sigma}_x^2$ to its variance σ_x^2, giving $\eta = |\widetilde{\sigma}_x^2|/\sigma_x^2$, which is bounded within [0, 1] [56], [57]. Then from (62), the SS-MSE $J(\infty)$ of CLMS can be derived as

$$\begin{aligned} J(\infty) &= \sigma_q^2 + \frac{\dfrac{\mu N\sigma_q^2\sigma_x^4}{2\sigma_x^2 - \mu\sigma_x^4 - \mu|\widetilde{\sigma}_x^2|^2}}{1 - \dfrac{\mu N\sigma_x^4}{2\sigma_x^2 - \mu\sigma_x^4 - \mu|\widetilde{\sigma}_x^2|^2}} \\ &= \sigma_q^2 + \frac{\mu N\sigma_q^2\sigma_x^2}{2 - \mu(N+1)\sigma_x^2 - \mu\sigma_x^2\eta^2} \end{aligned} \tag{65}$$

In a similar way, from (61), the SS-CMSE, $\widetilde{J}(\infty)$ of CLMS becomes

$$\begin{aligned} \widetilde{J}(\infty) &= \widetilde{\sigma}_q^2 + \frac{\dfrac{\mu N\widetilde{\sigma}_q^2|\widetilde{\sigma}_x^2|^2}{2\sigma_x^2 - 2\mu\sigma_x^4}}{1 - \dfrac{\mu N|\widetilde{\sigma}_x^2|^2}{2\sigma_x^2 - 2\mu\sigma_x^4}} \\ &= \widetilde{\sigma}_q^2 + \frac{\mu N\widetilde{\sigma}_q^2\sigma_x^2\eta^2}{2 - 2\mu\sigma_x^2 - \mu N\sigma_x^2\eta^2} \end{aligned} \tag{66}$$

Remark 5: Both the SS-MSE, $J(\infty)$, in (65) and the SS-CMSE, $\widetilde{J}(\infty)$, in (66) which is evaluated through its real and imaginary parts denoted as $\Re[\widetilde{J}(\infty)]$ and $\Im[\widetilde{J}(\infty)]$, are monotonically increasing functions of the degree of input noncircularity η. This is readily verified through the first derivative of $J(\infty)$, $\Re[\widetilde{J}(\infty)]$ and $\Im[\widetilde{J}(\infty)]$ with respect to η, which gives $\partial J(\infty)/\partial\eta > 0$, $\partial\Re[\widetilde{J}(\infty)]/\partial\eta > 0$ and $\partial\Im[\widetilde{J}(\infty)]/\partial\eta > 0$, for $\eta \in [0, 1]$.

V. Further Insights Into the Proposed Full Mean Square Analysis of CLMS

In general, the proposed complementary mean square analysis enables us to quantify the influence on the performance of the time-variant second order noncircularity of both the error $e(k)$ and the weight error vector $\widetilde{\mathbf{h}}(k)$ of CLMS, in both the transient and steady state stages, which is overlooked by the existing mean square analyses in the literature. Due to the complex-valued nature of the complementary second order statistics, we have two degrees of freedom to describe the noncircularity on $e(k)$ and $\widetilde{\mathbf{h}}(k)$, the two key parameters which govern the CLMS algorithm. This will become clear through the expressions for the individual mean square evolutions of the real and imaginary parts of $e(k)$ and $\widetilde{\mathbf{h}}(k)$, as follows. When $e(k)$ is rewritten in terms of its real and imaginary components as $e(k) = e_r(k) + \imath e_\imath(k)$, the complementary MSE $\widetilde{J}(k)$ in (19) becomes

$$\widetilde{J}(k) = E[e^2(k)] = E[e_r^2(k)] - E[e_\imath^2(k)] + 2\imath E[e_r(k)e_\imath(k)] \tag{67}$$

the real part of which represents the noncircularity (impropriety) introduced by the power mismatch between the real and imaginary components of the error $e(k)$, while the imaginary part of $\widetilde{J}(k)$ is twice of the cross-correlation between the two components of $e(k)$.

The complementary weight error covariance matrix $\widetilde{\mathbf{K}}(k)$ in (21) can be factorised in a similar way by considering $\widetilde{\mathbf{h}}(k) = \widetilde{\mathbf{h}}_r(k) + \imath\widetilde{\mathbf{h}}_\imath(k)$, to give

$$\begin{aligned} \widetilde{\mathbf{K}}(k) =\ & E[\widetilde{\mathbf{h}}(k)\widetilde{\mathbf{h}}^T(k)] \\ =\ & E[\widetilde{\mathbf{h}}_r(k)\widetilde{\mathbf{h}}_r^T(k)] - E[\widetilde{\mathbf{h}}_\imath(k)\widetilde{\mathbf{h}}_\imath^T(k)] \\ & + \imath\big(E[\widetilde{\mathbf{h}}_\imath(k)\widetilde{\mathbf{h}}_r^T(k)] + E[\widetilde{\mathbf{h}}_r(k)\widetilde{\mathbf{h}}_\imath^T(k)]\big) \end{aligned} \tag{68}$$

Note that the diagonal elements of $\widetilde{\mathbf{K}}(k)$ are of particular interest, since they quantify the time-varying second order noncircular behaviour of the weight error coefficients $\widetilde{\mathbf{h}}(k)$, in the sense that their real components reflects the power imbalance between the real part and imaginary part of $\widetilde{\mathbf{h}}(k)$, while the imaginary components of the diagonal elements in $\widetilde{\mathbf{K}}(k)$ are twice the cross-correlation between the real and imaginary components of the weight error coefficients $\widetilde{\mathbf{h}}(k)$.

In a similar way, we can partition $J(k)$ in (14) and $\mathbf{K}(k)$ in (16) to yield

$$\begin{aligned} J(k) &= E[|e(k)|^2] \\ &= E[e_r^2(k)] + E[e_\imath^2(k)] \end{aligned} \tag{69}$$

$$\begin{aligned} \mathbf{K}(k) &= E[\widetilde{\mathbf{h}}(k)\widetilde{\mathbf{h}}^H(k)] \\ &= E[\widetilde{\mathbf{h}}_r(k)\widetilde{\mathbf{h}}_r^T(k)] + E[\widetilde{\mathbf{h}}_\imath(k)\widetilde{\mathbf{h}}_\imath^T(k)] \\ &\quad + \imath\big(E[\widetilde{\mathbf{h}}_\imath(k)\widetilde{\mathbf{h}}_r^T(k)] - E[\widetilde{\mathbf{h}}_r(k)\widetilde{\mathbf{h}}_\imath^T(k)]\big) \end{aligned} \tag{70}$$

Upon inspection of both sides of (67) and (69), we obtain

$$E[e_r^2(k)] = \frac{J(k) + \Re[\widetilde{J}(k)]}{2} \tag{71}$$

$$E[e_\imath^2(k)] = \frac{J(k) - \Re[\widetilde{J}(k)]}{2} \tag{72}$$

$$E[e_r(k)e_\imath(k)] = \frac{\Im[\widetilde{J}(k)]}{2} \tag{73}$$

while a comparison of both sides of (68) and (70) yields

$$E[\widetilde{\mathbf{h}}_r(k)\widetilde{\mathbf{h}}_r^T(k)] = \frac{\Re[\mathbf{K}(k) + \widetilde{\mathbf{K}}(k)]}{2} \tag{74}$$

$$E[\widetilde{\mathbf{h}}_\imath(k)\widetilde{\mathbf{h}}_\imath^T(k)] = \frac{\Re[\mathbf{K}(k) - \widetilde{\mathbf{K}}(k)]}{2} \tag{75}$$

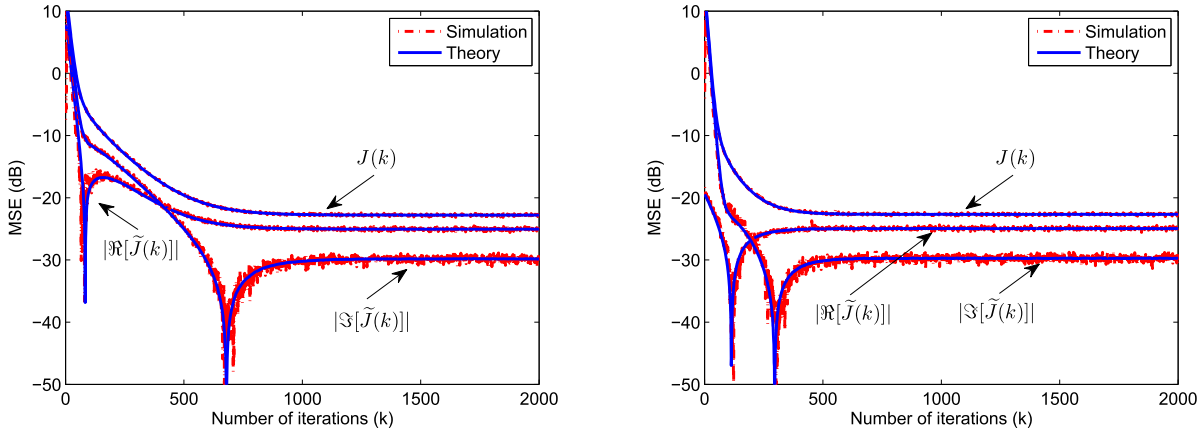


Fig. 3. Comparison of the theoretical and simulated curves for the MSE $J(k)$ and the CMSE $\widetilde{J}(k)$, quantified in terms of its real component $\Re[\widetilde{J}(k)]$ and imaginary component $\Im[\widetilde{J}(k)]$, with $\sigma_q^2 = 0.005$ and different step-sizes μ.

$$E[\widetilde{\mathbf{h}}_r(k)\widetilde{\mathbf{h}}_\imath^T(k)] = \frac{\Im[\widetilde{\mathbf{K}}(k) - \mathbf{K}(k)]}{2} \tag{76}$$

$$E[\widetilde{\mathbf{h}}_\imath(k)\widetilde{\mathbf{h}}_r^T(k)] = \frac{\Im[\widetilde{\mathbf{K}}(k) + \mathbf{K}(k)]}{2} \tag{77}$$

From (71)–(77), an important application of the proposed full mean square analysis of CLMS for second order noncircular Gaussian inputs stems from the fact that this provides enough degrees of freedom to describe the individual mean square evolutions of the real and imaginary parts, as well as their cross-correlation, of both the output error $e(k)$ and the weight error vector $\widetilde{\mathbf{h}}(k)$. This can only be achieved via a joint consideration of the proposed complementary and the standard mean square analyses together, when $e(k)$ and $\widetilde{\mathbf{h}}(k)$ are second order noncircular. The standard mean square analysis itself is adequate for this second order statistical decoupling only in a limited scenario when $e(k)$ and $\widetilde{\mathbf{h}}(k)$ are circular, i.e., $E[e_r^2(k)] = E[e_\imath^2(k)] = E[e^2(k)]/2$, $E[e_r(k)e_\imath(k)] = 0$, $E[\widetilde{\mathbf{h}}_r(k)\widetilde{\mathbf{h}}_r^T(k)] = E[\widetilde{\mathbf{h}}_\imath(k)\widetilde{\mathbf{h}}_\imath^T(k)] = E[\widetilde{\mathbf{h}}(k)\widetilde{\mathbf{h}}^T(k)]/2$, and $E[\widetilde{\mathbf{h}}_r(k)\widetilde{\mathbf{h}}_\imath^T(k)] = E[\widetilde{\mathbf{h}}_\imath(k)\widetilde{\mathbf{h}}_r^T(k)] = \mathbf{0}$. From this perspective, the proposed full mean square analysis is not limited to the CLMS algorithm, but is also applicable to other gradient descent based adaptive filtering algorithms, which may derived from different cost functions, e.g., Gaussian entropy criteria [36], [58], with second order noncircular inputs and system noises. The noncircularity of the output error and the weight error vector has also been observed within these algorithms.

VI. Simulations

Numerical examples were conducted in the MATLAB programming environment in order to evaluate the theoretical findings on the proposed complementary and full mean square convergence analysis of the CLMS algorithm for strictly linear estimation (SLE) with second order noncircular Gaussian input data. The experiments were performed in a system identification setting as described in (3), where the system coefficients to be identified, $\mathbf{h}^o$, formed a strictly linear FIR channel of length $N = 5$, for which the weight coefficients were drawn from a uniform complex-valued random distribution. The system noise $q(k)$ was a zero-mean complex-valued Gaussian process with variance σ_q^2, and with correlated real and imaginary parts obtained through

$$q_r(k) = \rho q_\imath(k) + \sqrt{1-\rho^2}\, q_1(k) \tag{78}$$

where ρ is a parameter which controls the correlation level between the two channels, set to be $\rho = 0.5$ in the simulations, and $q_1(k)$ is a real-valued Gaussian process with variance σ_q^2. In this way, the noncircularity of $q(k)$ arose from two sources, that is, from the power mismatch between the real and imaginary components and their cross-correlation. The Gaussian input $\mathbf{x}(k)$ was a second order noncircular linear autoregressive (AR)(1) process, given by

$$x(k) = 0.7x(k-1) + u(k) \tag{79}$$

where $u(k)$ is a zero-mean doubly white (DW) noncircular Gaussian process with variance $\sigma_u^2 = 1$ and complementary variance $\widetilde{\sigma}_u^2 = 0.9$, which gave a high degree of noncircularity $\eta = |\widetilde{\sigma}_u^2|/\sigma_u^2 = 0.9$. The weights within the CLMS algorithm were initialised with zeros, and simulation results were obtained by averaging over 10,000 independent trials.

Fig. 3 shows the theoretical and simulated learning curves of the standard MSE, $J(k)$, and the proposed complementary MSE (CMSE), $\widetilde{J}(k)$, of CLMS for the considered second order noncircular Gaussian input vector $\mathbf{x}(k)$ and the system noise $q(k)$ with $\sigma_q^2 = 0.005$ and different values of step-size μ. The theoretical evaluations of $J(k)$ and $\widetilde{J}(k)$, obtained by using (15) and (17), and (20) and (22), respectively, and their high accuracy in predicting the simulated behaviours can be observed. Note that in order to illustrate the convergence of both $J(k)$ and $\widetilde{J}(k)$ in detail, their logarithmic values in dB were plotted, and since the real and imaginary components of $\widetilde{J}(k)$, that is, $\Re[\widetilde{J}(k)]$ and $\Im[\widetilde{J}(k)]$ can be negative, their absolute values $|\Re[\widetilde{J}(k)]|$ and $|\Im[\widetilde{J}(k)]|$ were used. The simulation results are in line with *Remark 2*, which states that the noncircularity of both $\mathbf{x}(k)$ and

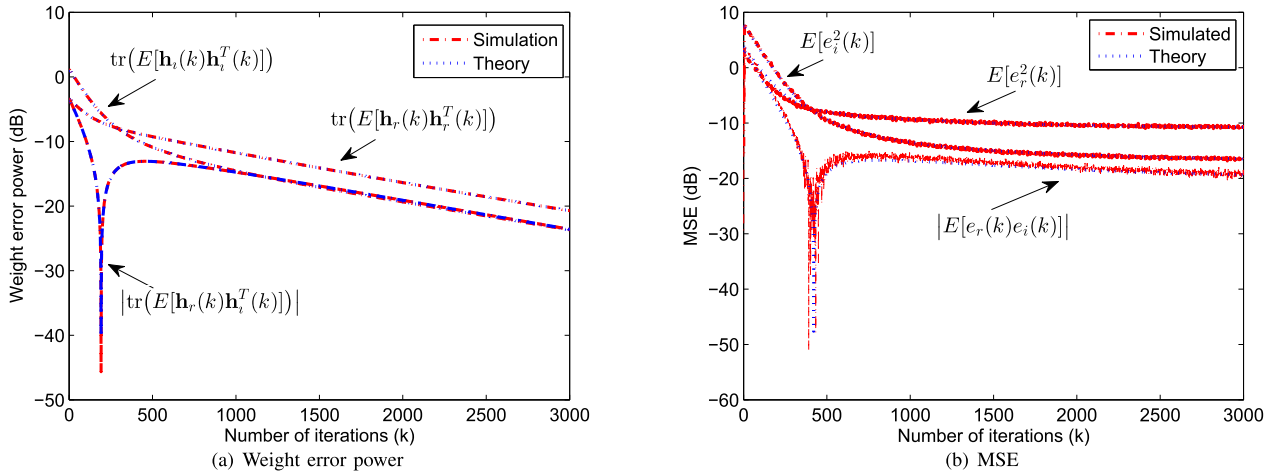


Fig. 4. Comparison of the theoretical and simulated curves for the mean square convergence of the real and imaginary parts of $e(k)$ and $\widetilde{\mathbf{h}}(k)$, as well as their cross-terms, with $\mu = 0.001$ and $\sigma_q^2 = 0.1$.

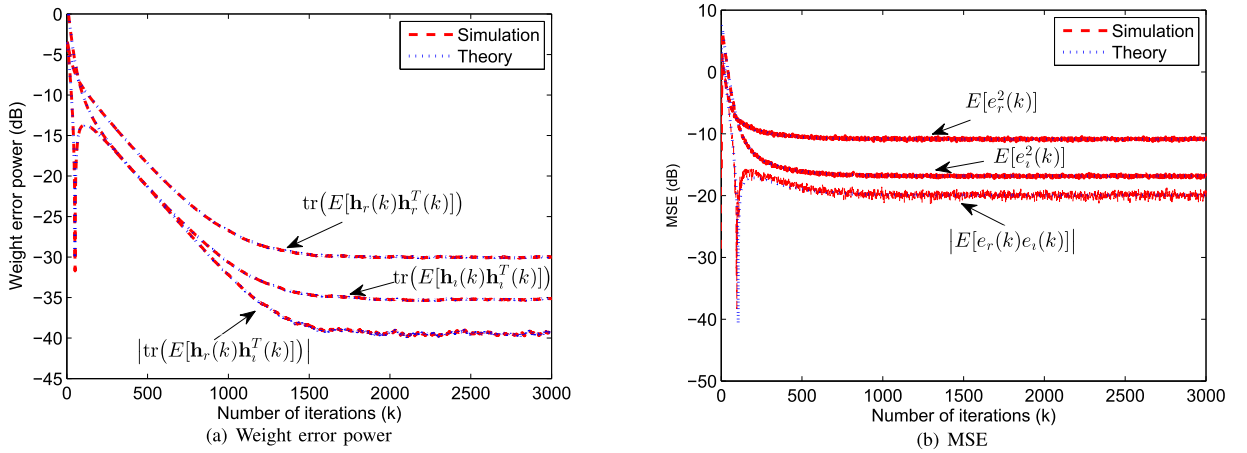


Fig. 5. Comparison of the theoretical and simulated curves for the mean square convergence of the real and imaginary parts of $e(k)$ and $\widetilde{\mathbf{h}}(k)$, as well as their cross-terms, with $\mu = 0.01$ and $\sigma_q^2 = 0.1$.

$q(k)$ would propagate into the output error $e(k)$, as indicated by a non-vanishing CMSE $\widetilde{J}(k)$. Furthermore, as discussed in *Remark 3*, the real and imaginary components of $\widetilde{J}(k)$ are both convergent during the recursive minimisation process of $J(k)$. Observe that there is a deep V-shape in the beginning of the evolutions of $|\Re[\widetilde{J}(k)]|$ and $|\Im[\widetilde{J}(k)]|$, because of a transition from a negative value to positive value.

As discussed in Section V, only the joint consideration of the proposed complementary mean square convergence analysis in (20) and (23), and the standard mean square convergence analysis in (15) and (18), enables a full insight into the evolution of the mean square error and the weight error covariance matrix of the real and imaginary channels, as well as their cross-terms, by using (71)–(77). This is supported by Fig. 4 and Fig. 5, where the results were obtained with $\sigma_q^2 = 0.1$ and for different μ. Note that in these simulations, for simplicity of illustration, we used the power of weight error coefficients, which is the trace of the weight error covariance matrix, instead of the whole matrix. It can be also seen from Fig. 4 and Fig. 5 that the theoretical analysis accurately described the full mean square convergence

TABLE I
COMPARISON OF THE EIGENVALUES OBTAINED BY EVD AND AUT

	λ_1	λ_2	λ_3	λ_4	λ_5
EVD	5.83984	1.94889	0.85500	0.49951	0.37274
AUT	5.83979	1.94890	0.85498	0.49952	0.37279

of the CLMS algorithm. Due to the fact that those cross-terms, that is, $E[e_r(k)e_i(k)]$ and $\mathrm{tr}\big(E[\widetilde{\mathbf{h}}_r(k)\widetilde{\mathbf{h}}_i^T(k)]\big)$, can be negative, we used their absolute values for the illustration. Note that the noncircularity of both $e(k)$ and $\widetilde{\mathbf{h}}(k)$ was also indicated by Fig. 4 and Fig. 5 in the sense that both the power mismatch between their real and imaginary components and the cross-correlation between both channels did exist.

In the next stage, we investigated the validity of the proposed full steady state performance analysis of CLMS. In order to achieve the closed-form expressions of the steady state MSE (SS-MSE), $J(\infty)$, and the steady state complementary MSE (SS-CMSE), $\widetilde{J}(\infty)$, of CLMS given in (62) and (61), the approximate uncorrelating transform (AUT) [34], [44] was employed to approximately diagonalise the input covariance matrix $\mathbf{R}$ by

TABLE II
COMPARISON OF THEORETICAL AND SIMULATED STEADY STATE MSE $J(\infty)$ AND CMSE $\widetilde{J}(\infty)$ OF THE CLMS ALGORITHM FOR SECOND ORDER NONCIRCULAR GAUSSIAN INPUT DATA AGAINST DIFFERENT VALUES OF THE STEP-SIZE μ AND THE SYSTEM NOISE VARIANCE σ_q^2

μ	σ_q^2	Theoretical (dB)			Simulated (dB)		
		$J(\infty)$ (62)	$\Re[\widetilde{J}(\infty)]$ (61)	$\Im[\widetilde{J}(\infty)]$ (61)	$J(\infty)$	$\Re[\widetilde{J}(\infty)]$	$\Im[\widetilde{J}(\infty)]$
0.001	0.001	-30.006	-32.226	-36.997	-29.945	-32.187	-36.944
0.005	0.001	-30.027	-32.256	-37.027	-29.923	-32.166	-36.906
0.005	0.01	-20.027	-22.256	-27.027	-19.949	-22.179	-26.944
0.01	0.05	-13.065	-15.304	-20.075	-12.883	-15.128	-19.896

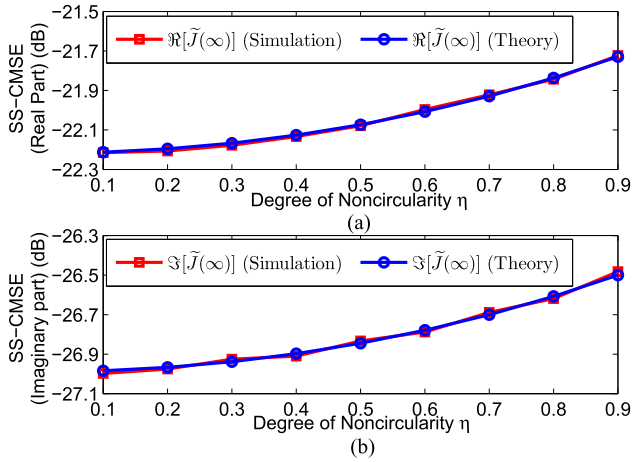


Fig. 6. Theoretical and simulated SS-CMSE, $\widetilde{J}(\infty)$, of CLMS against varying degrees of noncircularity η, for doubly white Gaussian input data with $\sigma_q^2 = 0.01$ and $\mu = 0.05$. (a) $\Re[\widetilde{J}(\infty)]$, and (b) $\Im[\widetilde{J}(\infty)]$.

using the same orthogonal matrix $\mathbf{Q}$ from the Takagi factorisation of its complementary counterpart $\mathbf{P}$, as given in (26). We should mention that for a highly noncircular univariate input $x(k)$, AUT gives a near perfect approximate diagonalisation for small sizes of covariance matrices, e.g., $N \leq 20$ [44]. This is evidenced by Table I, which compares the eigenvalues obtained by using the standard eigenvalue decomposition (EVD) and the AUT, and observe very tiny difference along all the five eigenvalues. Table II compares theoretical and simulated SS-MSE and SS-CMSE against different system noise variances σ_q^2 and step-sizes μ, showing a good agreement between the analytical and empirical results.

As discussed in Section IV-A, we have been able to build up an intuitive and explicit link between the full steady state MSE performance of both the CLMS algorithm and the degree of noncircularity η of the doubly white Gaussian input data. We have also shown that the steady state complementary MSE (SS-CMSE), $\widetilde{J}(\infty)$, in (66), expressed in terms of its real and imaginary components, monotonically increases with an increase in the input noncircularity η. To further illustrate this phenomenon, in the final experiment we evaluated the SS-CMSE, $\widetilde{J}(\infty)$, for varying degrees of input noncircularity η. Since the doubly white Gaussian input with unit variance was used in the simulations, for a fair comparison, different degrees of input noncircularity η were achieved by varying its complementary covariance $\widetilde{\sigma}_x^2$. The upper panel and the lower panel of Fig. 6 conform with the analysis in *Remark 5*, and the agreement between the simulated and theoretical SS-CMSE for the CLMS algorithm can also be observed in Fig. 6.

VII. CONCLUSION

A full mean square transient and steady state performance analysis of complex least mean square (CLMS) has been introduced for strictly linear estimation of second order noncircular Gaussian input data. This has been achieved by jointly considering the convergence of the complementary second order statistics of both the error and the weight error vector, namely the complementary mean square error (CMSE) and the complementary covariance matrix of the weight error vector, of the CLMS algorithm, together with their standard second order counterparts. We have also illustrated that there exists a conservative stability bound on the step-size for the convergence of both the weight error covariance matrix as its complementary counterpart. The closed-form expressions for the steady state MSE and CMSE have been subsequently derived, enabling us to find a monotonically increasing relationship between these quantities and the degree of input noncircularity for doubly white Gaussian input data. The proposed full mean square analysis has also provided us with enough degrees of freedom to analyse the mean square evolutions of the real and imaginary components of the error and the weight error vector independently, as well as the cross-correlation between both components, an important property of CLMS which cannot be achieved by using the standard mean square analysis only. Simulations in the system identification setting support the analysis.

APPENDIX A
THE STABILITY ANALYSIS OF THE TRANSITION MATRIX $\mathbf{F}$

Consider the eigenvalues of the transition matrix $\mathbf{F}$ in (31) as the solutions of an equation in β, that is

$$\det[\mathbf{F} - \beta\mathbf{I}] = 0 \tag{80}$$

By defining

$$\delta_i = 1 - 2\mu\lambda_i + \mu^2\lambda_i^2 + \mu^2 p_i^2, \ i = 1, 2, \ldots, N \tag{81}$$

the matrix $\mathbf{F}$ can be expressed as

$$\mathbf{F} = \mathrm{diag}\{\delta_1, \delta_2, \ldots, \delta_N\} + \mu^2 \boldsymbol{\lambda}_\mathbf{r} \boldsymbol{\lambda}_\mathbf{r}^T \tag{82}$$

and hence, after a few mathematical manipulations, we have

$$\det[\mathbf{F} - \beta\mathbf{I}] = \left[\prod_{i=1}^{N}(\delta_i - \beta)\right] \cdot \left[1 + \sum_{i=1}^{N} \frac{\mu^2\lambda_i^2}{\delta_i - \beta}\right] = 0 \tag{83}$$

which is equivalent to the equation

$$a(\beta) = 1 + \sum_{i=1}^{N} \frac{\mu^2 \lambda_i^2}{\delta_i - \beta} = 0 \tag{84}$$

whose poles are δ_i, all are positive and arranged in an increasing order, i.e., $0 < \delta_1 < \delta_2 < \cdots < \delta_N$. Since $\partial a(\beta)/\partial\beta > 0$, similar to the analysis of $f(\alpha)$ in (37), between each pair of successive poles of $a(\beta)$, there exists a single zero, so that

$$0 < \delta_1 < \beta_1 < \delta_2 < \beta_2 < \cdots < \beta_{N-1} < \delta_N < \beta_N \tag{85}$$

where β_i are the zeros of $a(\beta)$, and also the eigenvalues of $\mathbf{F}$. The stability condition that all the eigenvalues of $\mathbf{F}$ are located within the unit circle requires $\beta_N < 1$. This is the case if and only if

$$\delta_i < 1, \ i = 1, 2, \ldots, N \tag{86}$$

and

$$a(\beta) \mid_{\beta=1} = 1 + \sum_{i=1}^{N} \frac{\mu^2 \lambda_i^2}{\delta_i - \beta} > 0 \tag{87}$$

Substituting δ_i in (81) into (86) and (87) yields the following necessary and sufficient conditions

$$0 < \mu < \frac{2\lambda_i}{\lambda_i^2 + p_i^2}, \ i = 1, 2, \ldots, N \tag{88}$$

and

$$\sum_{i=1}^{N} \frac{\mu \lambda_i^2}{2\lambda_i - \mu(\lambda_i^2 + p_i^2)} - 1 < 0 \tag{89}$$

Define the left hand side of (89) by $b(\mu)$, then it will be a monotonically increasing function of μ with poles at $2\lambda_i/(\lambda_i^2 + p_i^2)$, ordered as $\lambda_1/(\lambda_1^2 + p_1^2) < \lambda_2/(\lambda_2^2 + p_2^2) < \cdots < \lambda_N/(\lambda_N^2 + p_N^2)$. Let μ_i be the solutions of $b(\mu) = 0$, then

$$0 < \mu_1 < \frac{2\lambda_N}{\lambda_N^2 + p_N^2} < \mu_2 < \frac{2\lambda_{N-1}}{\lambda_{N-1}^2 + p_{N-1}^2} < \cdots < \mu_N < \frac{2\lambda_1}{\lambda_1^2 + p_1^2}$$

The condition $0 < \mu < \mu_1$ now is equivalent to conditions in (88) and (89). Moreover, the polynomial $b(\mu) = 0$ can be rewritten as

$$\left(\frac{1}{\mu}\right)^N - m_1 \left(\frac{1}{\mu}\right)^{N-1} + m_2 \left(\frac{1}{\mu}\right)^{N-2} + \cdots + (-1)^N m_N = \prod_{i=1}^{N} \left(\frac{1}{\mu} - \frac{1}{\mu_i}\right) = 0 \tag{90}$$

where

$$m_1 = \frac{1}{2} \sum_{i=1}^{N} \left(2\lambda_i + \frac{p_i^2}{\lambda_i}\right) \tag{91}$$

$$m_2 = \frac{1}{4} \sum_{i=1}^{N} \sum_{j>i}^{N} \left(2\lambda_i + \frac{\lambda_i^2 + p_i^2}{\lambda_i}\right) \left(\frac{\lambda_j^2 + p_j^2}{\lambda_j}\right) \tag{92}$$

Since m_1 in (91) is equal to l_1 in (49), the equation (54) is also the conservative bound on μ to guarantee the convergence of $\boldsymbol{\kappa}(k)$ in (31).

REFERENCES

[1] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proc. IEEE*, vol. 63, no. 4, pp. 719–720, Apr. 1975.

[2] L. L. Horowitz and K. D. Senne, "Performance advantage of complex LMS for controlling narrow-band adaptive arrays," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-29, no. 3, pp. 722–736, Jun. 1981.

[3] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[4] B. Picinbono, "On circularity," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.

[5] B. Fisher and N. J. Bershad, "The complex LMS adaptive algorithm−Transient weight mean and covariance with applications to the ALE," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-31, no. 1, pp. 34–44, Feb. 1983.

[6] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. Signal Process.*, vol. 41, no. 9, pp. 2811–2825, Sep. 1993.

[7] M. Rupp and A. H. Sayed, "A time-domain feedback analysis of filtered-error adaptive gradient algorithms," *IEEE Trans. Signal Process.*, vol. 44, no. 6, pp. 1428–1439, Jun. 1996.

[8] L. Lindbom, M. Sternad, and A. Ahlén, "Tracking of time-varying mobile radio channels−Part I: The Wiener LMS algorithm," *IEEE Trans. Commun*, vol. 49, no. 12, pp. 2207–2217, Dec. 2001.

[9] A. Ahlén, L. Lindbom, and M. Sternad, "Analysis of stability and performance of adaptation algorithms with time-invariant gains," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 103–116, Jan. 2004.

[10] M. Godavarti and A. O. Hero III, "Partial update LMS algorithms," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2382–2399, Jul. 2005.

[11] Y. Avargel and I. Cohen, "Performance analysis of cross-band adaptation for subband acoustic echo cancellation," in *Proc. Int. Workshop Acoust. Echo Noise Control*, Sep. 2006, pp. 1–4.

[12] Y. Avargel and I. Cohen, "Adaptive nonlinear system identification in the short-time Fourier transform domain," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3891–3904, Oct. 2009.

[13] X. Zhao and A. H. Sayed, "Performance limits for distributed estimation over LMS adaptive networks," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5107–5124, Oct. 2012.

[14] A. van den Bos, "The multivariate complex normal distribution−A generalization," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 537–539, Mar. 1995.

[15] B. Picinbono, "Second-order complex random vectors and normal distributions," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.

[16] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, Feb. 1997.

[17] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and process," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 714–725, Mar. 2003.

[18] D. P. Mandic and S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. New York, NY, USA: Wiley, 2009.

[19] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

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

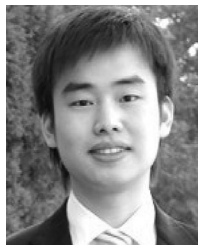
[21] W. H. Gerstacher, R. Schober, and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1512–1523, Sep. 2003.

[22] P. Chevalier and F. Pipon, "New insights into optimal widely linear array receivers for the demodulation of BPSK, MSK and GMSK signals corrupted by noncircular interferences—Application to SAIC," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 746–756, Mar. 2006.

[23] L. Anttila, M. Valkama, and M. Renfors, "Circularity-based I/Q imbalance compensation in wideband direct-conversion receivers," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2099–2113, Jul. 2008.

[24] D. P. Mandic, S. Javidi, S. L. Goh, A. Kuh, and K. Aihara, "Complex-valued prediction of wind profile using augmented complex statistics," *Renewable Energy*, vol. 34, no. 1, pp. 44–54, Jan. 2009.

[25] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5101–5123, Nov. 2011.
[26] Y. Xia, B. Jelfs, M. M. Van Hulle, J. C. Principe, and D. P. Mandic, "An augmented echo state network for nonlinear adaptive filtering of complex noncircular signals," *IEEE Trans. Neural Netw.*, vol. 22, no. 1, pp. 74–83, Jan. 2011.
[27] 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. 29, no. 5, pp. 44–54, Sep. 2012.
[28] 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.
[29] B. Jelfs, D. P. Mandic, and S. C. Douglas, "An adaptive approach for the identification of improper complex signals," *Signal Process.*, vol. 92, no. 2, pp. 335–344, 2012.
[30] T. Adali and P. J. Schreier, "Optimization and estimation of complex-valued signals: Theory and applications in filtering and blind source separation," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 112–128, Sep. 2014.
[31] Z. Li, Y. Xia, W. Pei, K. Wang, Y. Huang, and D. P. Mandic, "Noncircular measurement and mitigation of I/Q imbalance for OFDM-based WLAN transmitters," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 3, pp. 383–393, Mar. 2017.
[32] 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.
[33] D. P. Mandic, Y. Xia, and S. C. Douglas, "Steady state analysis of the CLMS and augmented CLMS algorithms for noncircular complex signals," in *Proc. 44th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2010, pp. 1635–1639.
[34] D. P. Mandic, S. Kanna, and S. C. Douglas, "Mean square analysis of the CLMS and ACLMS for non-circular signals: The approximate uncorrelating transform approach," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2015, pp. 3531–3535.
[35] S. C. Douglas and D. P. Mandic, "Mean and mean-square analysis of the complex LMS algorithm for non-circular Gaussian signals," in *Proc. IEEE 13th Digit. Signal Process. Workshop, 5th IEEE Signal Process. Educ. Workshop*, Apr. 2009, pp. 101–106.
[36] X.-L. Li and T. Adali, "Complex-valued linear and widely linear filtering using MSE and Gaussian entropy," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5672–5684, Nov. 2012.
[37] J. G. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 1995.
[38] P. Ciblat and L. Vandendorpe, "Blind carrier frequency offset estimation for noncircular constellation-based transmissions," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1378–1389, May 2003.
[39] T. Fusco and M. Tanda, "Blind frequency-offset estimation for OFDM/OQAM systems," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1828–1838, May 2007.
[40] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Transmit optimization with improper Gaussian signaling for interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2899–2913, Jun. 2013.
[41] Y. Zeng, R. Zhang, E. Gunawan, and Y. L. Guan, "Optimized transmission with improper Gaussian signaling in the *K*-user MISO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6303–6313, Dec. 2013.
[42] H. D. Nguyen, R. Zhang, and S. Sun, "Improper signaling for symbol error rate minimization in *K*-user MISO interference channel," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 857–869, Mar. 2015.
[43] Y. Xia, Z. Blazic, and D. P. Mandic, "Complex-valued least squares frequency estimation for unbalanced power systems," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 3, pp. 638–648, Mar. 2015.
[44] C. Cheong Took, S. C. Douglas, and D. P. Mandic, "On approximate diagonalization of correlation matrices in widely linear signal procesing," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1469–1473, Mar. 2012.
[45] A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 33, no. 1, pp. 222–230, Feb. 1985.
[46] S. C. Douglas, "Analysis of the multiple-error and block least-mean-square adaptive algorithms," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 42, no. 2, pp. 92–101, Feb. 1995.
[47] N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyzes of adaptive filters," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 314–324, Feb. 2001.
[48] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York, NY, USA: Wiley, 2003.
[49] J. E. Mazo, "On the independence theory of equalizer convergence," *Bell Syst. Tech. J.*, vol. 58, no. 3, pp. 963–993, May 1979.
[50] S. C. Douglas and W. Pan, "Exact expectation analysis of the LMS adaptive filter," *IEEE Trans. Signal Process.*, vol. 43, no. 12, pp. 2863–2871, Dec. 1995.
[51] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: Identifiability, uniqueness, and separability," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1017–1029, Mar. 2006.
[52] S. C. Douglas, "Fixed-point algorithms for the blind separation of arbitrary complex-valued non-Gaussian signal mixtures," *EURASIP J. Adv. Signal Process.*, vol. 2007, 2007, Art. no. 36525.
[53] R. A. Horn, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.
[54] O. L. Jagerman, "Nonstationary blocking in telephone traffic," *Bell Syst. Tech. J.*, vol. 54, no. 3, pp. 625–661, Mar. 1975.
[55] K. Mayyas, "Performance analysis of the deficient length LMS adaptive algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2727–2640, Aug. 2005.
[56] E. Ollila, "On the circularity of a complex random variable," *IEEE Signal Process. Lett.*, vol. 15, pp. 841–844, 2008.
[57] E. Ollila, J. Eriksson, and V. Koivunen, "Complex elliptically symmetric random variables−Generation, characterization, and circularity tests," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 58–69, Jan. 2011.
[58] S. Huang, C. Li, and Y. Liu, "Complex-valued filtering based on the minimization of complex-error entropy," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 5, pp. 695–708, May 2013.



Yili Xia (M'11) received the B.Eng. degree in information engineering from Southeast University, Nanjing, China, in 2006, and the M.Sc. (Distinction) degree in communications & signal processing from the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K., in 2007, and the Ph.D. degree in adaptive signal processing from Imperial College London, London, in 2011.

Since 2013, he has been an Associate Professor with the School of Information and Engineering, Southeast University, Nanjing, where he is currently the Deputy Head in the Department of Information and Signal Processing Engineering. His research interests include complex and hypercomplex statistical analysis, linear and nonlinear adaptive filters, as well as their applications on communications and power systems.

Danilo P. Mandic (M'99–SM'03–F'12) received the Ph.D. degree in nonlinear adaptive signal processing from Imperial College London, London, U.K., in 1999.

He is currently a Professor in signal processing in Imperial College London. He has been working in the area of nonlinear adaptive signal processing, multivariate data analysis, and nonlinear dynamics. He has been a Guest Professor at Katholieke Universiteit Leuven, Leuven, Belgium, Tokyo University of Agriculture and Technology, Tokyo, Japan, and Westminster University, London, and a Frontier Researcher at RIKEN, Japan. His publication record includes two research monographs entitled *Recurrent Neural Networks for Prediction*: *Learning Algorithms*, *Architectures and Stability* (1st ed., Aug. 2001) and *Complex Valued Nonlinear Adaptive Filters*: *Noncircularity*, *Widely Linear and Neural Models* (1st ed., Wiley, Apr. 2009), an edited book titled *Signal Processing Techniques for Knowledge Extraction and Information Fusion* (Springer, 2008), and more than 200 publications on signal and image processing.

Dr. Mandic has been a member of the IEEE Technical Committee on Signal Processing Theory and Methods, and has been an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON NEURAL NETWORKS, and the *International Journal of Mathematical Modelling and Algorithms*. He is currently an Associate Editor for the IEEE SIGNAL PROCESSING MAGAZINE and IEEE TRANSACTIONS ON INFORMATION AND SIGNAL PROCESSING OVER NETWORKS. He has produced award winning papers and products resulting from his collaboration with Industry.